

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:57:41 ; Search time 37.6351 Seconds  
(without alignments)  
3819.179 Million cell updates/sec

Title: US-09-647-978A-2

Perfect score: 2832

Sequence: 1 MSDGTASARSSPLDRPAF.....LSCAESEDPCRELTPDKS 557

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 830525

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- 1: sp\_archaea.\*
- 2: sp\_bacteria.\*
- 3: sp\_fungi.\*
- 4: sp\_human.\*
- 5: sp\_invertebrate.\*
- 6: sp\_mammal.\*
- 7: sp\_mhc.\*
- 8: sp\_organelle.\*
- 9: sp\_phage.\*
- 10: sp\_plant.\*
- 11: sp\_rodent.\*
- 12: sp\_virus.\*
- 13: sp\_vertebrate.\*
- 14: sp\_unclassified.\*
- 15: sp\_virus.\*
- 16: sp\_bacteriap.\*
- 17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	2832	100.0	557	11 Q9WV89	Q9WV89 mus musculus
2	2522	89.1	533	11 Q8CFL1	Q8CFL1 mus musculus
3	755	26.7	245	4 Q81V25	Q81V25 homo sapien
4	195.5	6.9	1492	4 Q9NU93	Q9NU93 homo sapien
5	192.5	6.8	2484	6 Q28006	Q28006 bos taurus
6	187.5	6.6	1427	4 Q9NU94	Q9NU94 homo sapien
7	183.5	6.5	1333	4 Q9NU91	Q9NU91 homo sapien
8	183	6.5	1829	11 Q35889	Q35889 rattus norv
9	182.5	6.4	883	4 Q96MN6	Q96MN6 homo sapien
10	181.5	6.4	1410	4 Q9NSN7	Q9NSN7 homo sapien
11	178	6.3	1296	4 Q9NU90	Q9NU90 homo sapien
12	177.5	6.3	1663	11 Q35890	Q35890 rattus norv
13	175	6.2	1624	5 Q9U679	Q9U679 strongyloce
14	174	6.1	1136	4 Q8TEW8	Q8TEW8 homo sapien
15	171.5	6.1	1205	4 Q96NX7	Q96NX7 homo sapien
16	171.5	6.1	1205	4 Q81UC7	Q81UC7 homo sapien

17	168	5.9	1104	4 Q81UC9	Q81UC9 homo sapien
18	167	5.9	1116	3 Q9HGL2	Q9HGL2 schizosacch
19	165	5.8	933	5 Q9W3V2	Q9W3V2 drosophila
20	165	5.8	2176	5 Q46112	Q46112 drosophila
21	165	5.8	2460	11 Q64512	Q64512 mus musculus
22	164.5	5.8	561	5 Q95Q64	Q95Q64 caenorhabdi
23	164.5	5.8	658	5 Q95Q66	Q95Q66 caenorhabdi
24	164.5	5.8	721	5 P91146	P91146 caenorhabdi
25	164	5.8	927	11 Q62402	Q62402 mus musculus
26	163.5	5.8	611	5 Q9V6P9	Q9V6P9 drosophila
27	163.5	5.8	1459	3 P87198	P87198 ustilago ma
28	163	5.8	411	4 Q8ENH7	Q8ENH7 homo sapien
29	163	5.8	624	4 Q96DK9	Q96DK9 homo sapien
30	162.5	5.7	963	4 Q96HNS	Q96HNS homo sapien
31	161.5	5.7	1220	13 Q8JFT5	Q8JFT5 brachydanio
32	161.5	5.7	1721	13 Q8JFT4	Q8JFT4 brachydanio
33	160.5	5.7	1006	4 Q96157	Q96157 homo sapien
34	160.5	5.7	1871	5 Q81R54	Q81R54 drosophila
35	160.5	5.7	1889	4 Q9H430	Q9H430 homo sapien
36	160.5	5.7	2010	4 Q9P216	Q9P216 homo sapien
37	160.5	5.7	2328	5 Q9VY43	Q9VY43 drosophila
38	160.5	5.7	2360	5 Q81R55	Q81R55 drosophila
39	160	5.6	888	5 Q81P21	Q81P21 drosophila
40	160	5.6	930	5 Q81P22	Q81P22 drosophila
41	160	5.6	1103	5 Q8MR15	Q8MR15 drosophila
42	160	5.6	1658	5 Q9BIC1	Q9BIC1 caenorhabdi
43	160	5.6	1666	5 Q9XY66	Q9XY66 caenorhabdi
44	160	5.6	2162	5 Q9VQM0	Q9VQM0 drosophila
45	159.5	5.6	574	11 Q8RID2	Q8RID2 mus musculus

ALIGNMENTS

RESULT 1

Q9WV89  
ID Q9WV89 PRELIMINARY; PRT; 557 AA.  
AC Q9WV89;  
DT 01-NOV-1999 (Tremblrel. 12, Created)  
DT 01-NOV-1999 (Tremblrel. 12, Last sequence update)  
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)  
DE Syntaxin4-interacting protein synip.  
GN STXBp4.  
OS Mus musculus (Mouse).  
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.  
OX NCBI\_TaxID=10090;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=99322657; PubMed=10394363;  
RA Min J., Okada S., Kanzaki M., Elmendorf J.S., Coker K.J., Ceresa B.P.,  
Syu L.J., Noda Y., Saltiel A.R., Pessin J.B.;  
RT "Synip: a novel insulin-regulated syntaxin 4-binding protein mediating  
RT GLUT4 translocation in adipocytes.";  
RL Mol. Cell 3:751-760(1999).  
CC -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.  
DR EMBL; AF152924; AAD43533.1; -;  
DR MGD; MGI:1342296; Stxbp4.  
DR InterPro; IPR001478; PDZ.  
DR InterPro; IPR001202; WW\_Rsp5\_WWP.  
DR Pfam; PF00595; PDZ; 1.  
DR Pfam; PF00397; WW; 1.  
DR SMART; SM00228; PDZ; 1.  
DR SMART; SM00456; WW; 1.  
DR PROSITE; PSS0106; PDZ; 1.  
DR PROSITE; PSS0159; WW\_DOMAIN\_1; 1.  
DR PROSITE; PSS0020; WW\_DOMAIN\_2; 1.  
SQ SEQUENCE 557 AA; 61688 MW; 9211A8B02AF8EC96 CRC64;

Query Match 100.0%; Score 2832; DB 11; Length 557;  
Best Local Similarity 100.0%; Pred. No. 1.1e-157;  
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

June 1997

```

QY 1 MSDGTASARSSPLDRDPAPFVITVTKETGLGLKILGINRNEGPLYIHEVTPGGDCYK 60
DB 1 MSDGTASARSSPLDRDPAPFVITVTKETGLGLKILGINRNEGPLYIHEVTPGGDCYK 60
QY 61 DGRLLPGDQLVSIKESMIGVSFFEAASIIITRAKLRSFESPWETAFIROKSYCGHGNICC 120
DB 61 DGRLLPGDQLVSIKESMIGVSFFEAASIIITRAKLRSFESPWETAFIROKSYCGHGNICC 120
QY 121 PSPOVSEDCGPTSTFTLLSSPSETLLPKTSSPTQDSTFPSCKAIOTKPEHDKTEHSP 180
DB 121 PSPOVSEDCGPTSTFTLLSSPSETLLPKTSSPTQDSTFPSCKAIOTKPEHDKTEHSP 180
QY 181 ITSLDNSPADTSNADIAPAWTDGSGPQGIKISLNPVRLKAEKLEMANLYLGIOPTKEOR 240
DB 181 ITSLDNSPADTSNADIAPAWTDGSGPQGIKISLNPVRLKAEKLEMANLYLGIOPTKEOR 240
QY 241 EALREQVOADSKGTVSFGDFVQVARSIFCLOLDEVNVGVHEIPSILDSQLLPCDSLEADE 300
DB 241 EALREQVOADSKGTVSFGDFVQVARSIFCLOLDEVNVGVHEIPSILDSQLLPCDSLEADE 300
QY 301 VGLKQERNAALERNVLKELLESSEKHKQKQIEELQNVKQEKAKAVAEETRALRSIHILA 360
DB 301 VGLKQERNAALERNVLKELLESSEKHKQKQIEELQNVKQEKAKAVAEETRALRSIHILA 360
QY 361 EAAQORAHGEMDMYEEVIRLLEAEVSELKAQADYSDONKESVQDLRKRVTVLDCQLRKS 420
DB 361 EAAQORAHGEMDMYEEVIRLLEAEVSELKAQADYSDONKESVQDLRKRVTVLDCQLRKS 420
QY 421 EMARKAFKASTERLLGFTEAQEVLLDSSAPLSTLSERRAVLASOTSLLPARNGRSFFA 480
DB 421 EMARKAFKASTERLLGFTEAQEVLLDSSAPLSTLSERRAVLASOTSLLPARNGRSFFA 480
QY 481 TLLLESKELVRSVRAILDMDC 501
DB 481 TLLLESKELVRSVRAILDMDC 501
QY 541 AEESEDCPRELTPDKS 557
DB 541 AEESEDCPRELTPDKS 557
RESULT 2
Q8CFL1 PRELIMINARY; PRT; 533 AA.
AC Q8CFL1;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DE Similar to syntaxin binding protein 4.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Thymus;
RA Strausberg R.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC022861; AAH32861.1; -.
SQ SEQUENCE 533 AA; 58591 MW; 886BC19511E1JAF09 CRC64;
Query Match 89.1%; Score 2522; DB 11; Length 533;
Best Local Similarity 100.0%; Pred. No. 1.4e-139;
Matches 501; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDGTASARSSPLDRDPAPFVITVTKETGLGLKILGINRNEGPLYIHEVTPGGDCYK 60
DB 1 MSDGTASARSSPLDRDPAPFVITVTKETGLGLKILGINRNEGPLYIHEVTPGGDCYK 60
QY 61 DGRLLPGDQLVSIKESMIGVSFFEAASIIITRAKLRSFESPWETAFIROKSYCGHGNICC 120
DB 61 DGRLLPGDQLVSIKESMIGVSFFEAASIIITRAKLRSFESPWETAFIROKSYCGHGNICC 120

```

```

QY 121 PSPOVSEDCGPTSTFTLLSSPSETLLPKTSSPTQDSTFPSCKAIOTKPEHDKTEHSP 180
DB 121 PSPOVSEDCGPTSTFTLLSSPSETLLPKTSSPTQDSTFPSCKAIOTKPEHDKTEHSP 180
QY 181 ITSLDNSPADTSNADIAPAWTDGSGPQGIKISLNPVRLKAEKLEMANLYLGIOPTKEOR 240
DB 181 ITSLDNSPADTSNADIAPAWTDGSGPQGIKISLNPVRLKAEKLEMANLYLGIOPTKEOR 240
QY 241 EALREQVOADSKGTVSFGDFVQVARSIFCLOLDEVNVGVHEIPSILDSQLLPCDSLEADE 300
DB 241 EALREQVOADSKGTVSFGDFVQVARSIFCLOLDEVNVGVHEIPSILDSQLLPCDSLEADE 300
QY 301 VGLKQERNAALERNVLKELLESSEKHKQKQIEELQNVKQEKAKAVAEETRALRSIHILA 360
DB 301 VGLKQERNAALERNVLKELLESSEKHKQKQIEELQNVKQEKAKAVAEETRALRSIHILA 360
QY 361 EAAQORAHGEMDMYEEVIRLLEAEVSELKAQADYSDONKESVQDLRKRVTVLDCQLRKS 420
DB 361 EAAQORAHGEMDMYEEVIRLLEAEVSELKAQADYSDONKESVQDLRKRVTVLDCQLRKS 420
QY 421 EMARKAFKASTERLLGFTEAQEVLLDSSAPLSTLSERRAVLASOTSLLPARNGRSFFA 480
DB 421 EMARKAFKASTERLLGFTEAQEVLLDSSAPLSTLSERRAVLASOTSLLPARNGRSFFA 480
QY 481 TLLLESKELVRSVRAILDMDC 501
DB 481 TLLLESKELVRSVRAILDMDC 501
RESULT 3
Q8IVZ5 PRELIMINARY; PRT; 245 AA.
AC Q8IVZ5;
DT 01-MAR-2003 (Tremblrel. 23, Created)
DT 01-MAR-2003 (Tremblrel. 23, Last sequence update)
DT 01-MAR-2003 (Tremblrel. 23, Last annotation update)
DE Similar to syntaxin binding protein 4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Eye;
RA Strausberg R.;
RL Submitted (DEC-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; BC041485; AAH41485.1; -.
SQ SEQUENCE 245 AA; 26974 MW; 4E92DDBBFD27DF79 CRC64;
Query Match 26.7%; Score 755; DB 4; Length 245;
Best Local Similarity 67.3%; Pred. No. 1e-36;
Matches 165; Conservative 23; Mismatches 49; Indels 8; Gaps 5;
QY 78 MIGVSFFEAASIIITRAKL--RSESPWEIAFIROKSYCGHFNICCPVS-POVSEDCGPTQS 134
DB 1 MIGVSFFEAASIIITRAKLSTLESAWEIAFIROKSDNIQENLSCTSLTEASGEYGPQAS 60
QY 135 TFTLLSSPSETLLPKTSSPTQDSTFPSCKAIOTKPEHDKTEHSPITSLDNSPADTSNA 194
DB 61 TSLFSSPPEILIPKTSSTPTKTNNDILSSC--EIKTYNKTVOIPIITS-ENSTVGLSNT 116
QY 195 DIAPAWTDGSGPQGIKISLNPVRLKAEKLEMANLYLGIOPTKEORAEAREQVQADSKGT 254
DB 117 DVASANT-ENYGLQEKISLNPVRFKAEKLEMANLYLGIOPTKEORQALRQVQADSKGT 175
QY 255 VSGDFVQVARSIFCLOLDEVNVGVHEIPSILDSQLLPCDSLEADEVGKLRQERNAALEE 314
DB 176 VSGDFVQVARNLFCLOLDEVNVGAHEISNILDQLPCDSSEADEMERKLCERDQALKE 235
QY 315 RNVLK 319
DB 236 VNTLK 240

```

QY	ID	Q28006	PRELIMINARY;	PRT;	1492 AA.
QY	310	AALEERNVLEKLLSEKRRK	-----QLIEELQNVKOEAKAVAEETALRSRI	-----	357
DB	1188	A-----EKQOMQHIVDMLSKEIQLOSKPDRSAEDRLKRLMLEWFOK	1236		
QY	358	HLAAQAQAQMGEMDYEV	---IRLLAEAVSELKAQLADYSQNKESVQDLKRKRVTLDC	415	
DB	1237	RLQESKQKDEDEEDDDVTMLMQLEAERRARLODEERRRQOOLEMKKEA--ED	1294		
QY	416	QURKSEMARKAFKASTERLLGFIEAIOEVLDDSSAPLSTISERRAVLASQTSPLPLARNG	475		
DB	1295	RARQEEERRRQEEETKR	-----DAEKKRQEEGYYSRLAEARRQ	1335	
QY	476	RSFPATLLSEKELVRSVRAILDMDCPLPGWEEAYTAD	-----GIKYFINHVTTQTSWI	529	
DB	1336	HDEAARRLLE-PEAPGLCRPPLPRDYEPSPSPAGAPPPPPQPNASLYLKTQVLSPOSILF	1394		
QY	530	HPVMSALNLSCAESERDCPRELTDPKS	557		
DB	1395	TAKFVAYN---EEEEEDC--SLAGPNS	1417		
QY	310	AALEERNVLEKLLSEKRRK	-----QLIEELQNVKOEAKAVAEETALRSRI	-----	357
DB	1188	A-----EKQOMQHIVDMLSKEIQLOSKPDRSAEDRLKRLMLEWFOK	1236		
QY	358	HLAAQAQAQMGEMDYEV	---IRLLAEAVSELKAQLADYSQNKESVQDLKRKRVTLDC	415	
DB	1237	RLQESKQKDEDEEDDDVTMLMQLEAERRARLODEERRRQOOLEMKKEA--ED	1294		
QY	416	QURKSEMARKAFKASTERLLGFIEAIOEVLDDSSAPLSTISERRAVLASQTSPLPLARNG	475		
DB	1295	RARQEEERRRQEEETKR	-----DAEKKRQEEGYYSRLAEARRQ	1335	
QY	476	RSFPATLLSEKELVRSVRAILDMDCPLPGWEEAYTAD	-----GIKYFINHVTTQTSWI	529	
DB	1336	HDEAARRLLE-PEAPGLCRPPLPRDYEPSPSPAGAPPPPPQPNASLYLKTQVLSPOSILF	1394		
QY	530	HPVMSALNLSCAESERDCPRELTDPKS	557		
DB	1395	TAKFVAYN---EEEEEDC--SLAGPNS	1417		

RESULT 5

Q28006 PRELIMINARY; PRT: 2484 AA.

ID Q28006 AC Q28006; AC Q28006; DT 01-NOV-1996 (TremBLrel. 01, Created) DT 01-NOV-1996 (TremBLrel. 01, Last sequence update) DT 01-OCT-2002 (TremBLrel. 22, Last annotation update) DE BA14 tyrosine phosphatase (EC 3.1.3.48). OS Bos taurus (Bovine) OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; OC Mammalia; Eutheria; Cetartiodactyla; Ruminantia; Pecora; Bovidae; OC Bovidae; Bovinae; Bos. OX NCBI\_TaxID=9913; RN [1] RP SEQUENCE FROM N.A. RA Vega Q.C., Walton K.M., Dixon J.E.; RL Submitted (FEB-1995) to the EMBL/GenBank/DBJ databases. CC -1- CATALYTIC ACTIVITY: PROTEIN TYROSINE PHOSPHATE + H(2)O = PROTEIN TYROSINE + PHOSPHATE. CC -1- SIMILARITY: CONTAINS A DOMAIN FOUND IN BAND 4.1, EZRIN, MOESIN, CC -1- SIMILARITY: BELONGS TO THE NON-RECEPTOR CLASS OF THE PROTEIN-TYROSINE PHOSPHATASE FAMILY. ENBL: U20807; AAA73516.1; HSSP: Q12923; 3P0Z. DR InterPro: IPR000299; Band\_4.1. DR InterPro: IPR001478; PDZ. DR InterPro: IPR000387; TYR phosphatase. DR InterPro: IPR000242; TYR PP. DR Pfam: PF00373; Band\_41; 1. DR Pfam: PF00595; PDZ\_5. DR Pfam: PF00102; Y\_phosphatase; 1. DR PRINTS; PRO00935; BAND41. DR PRINTS; PRO0700; PRTPHPTASE. DR SMART; SM00295; B41; 1. DR SMART; SM00228; PDZ; 5. DR SMART; SM00194; PTPC; 1. DR PROSITE; PS00660; BAND\_41\_1; FALSE NEG. DR PROSITE; PS00661; BAND\_41\_2; FALSE NEG. DR PROSITE; PS00567; BAND\_41\_3; 1. DR PROSITE; PS00106; PDZ; 5. DR PROSITE; PS00383; TYR PHOSPHATASE 1; FALSE\_NEG. DR PROSITE; PS00566; TYR PHOSPHATASE 2; 1. DR PROSITE; PS00555; TYR PHOSPHATASE\_PTP; 1. KW Structural protein; Cytoskeleton; Hydrolase; Coiled coil. FT DOMAIN 600 800 BAND 4.1-LIKE DOMAIN. FT DOMAIN 1341 1344 POLY-SER. FT DOMAIN 2236 2484 PROTEIN-TYROSINE PHOSPHATASE. FT ACT\_SITE 2407 2407 BY SIMILARITY. FT DOMAIN 467 504 COILED COIL (POTENTIAL).

Query Match 6.9%; Score 195.5; DB 4; Length 1492; Best Local Similarity 18.8%; Pred. No. 0.0049; Matches 152; Conservative 103; Mismatches 236; Indels 317; Gaps 31;

QY 11 SSPLDRPAPRVITVKTETGLGKILGGINRNEGVL-VYIHEVIGGDCYKDGRLKPGDQ 69

DB 666 AQLRKEPEITVTLKQNGMGLSIVAAGAGODKIGIYVKSXGGAADVDRLAGDQ 725

QY 70 LVSINKESMIGVSFEAKSIITRAKLRESPEWIAFIQKS-YCGHGNICCPSP---QV 125

DB 726 LLSVDGRSLVGLSQRAAELMRTS--SVVTLLEVA--KQAIYHGLATLLNQSPHMQRI 781

QY 126 SEDCG-----PQTSFTLLSSPETLLPKTSSTPQDTS----- 159

DB 782 SDRGSGKPRKSGEGPELYNNSTQNGSPSPOLPMAEYSEPKLPQDRLMKQRADHRSS 841

QY 160 -----TF 161

DB 842 PNVAQPPSPGKSAVASGTTAKITSVSTGNLCTEBQTPPRPEAYPIPTQTYREYTF 901

QY 162 PSCKA-----LQTKPE-HDKTEHSP 181

DB 902 PASKSQDMAPPQNWPNVEEKPMPHTDNLHNSIAIQVTRSQBELREDKAYQLERHRIE 961

QY 182 -----TSLDNSPA-----DTSNADIAPAWTDDSGPOCK 210

DB 962 AAMDRKSDSDMINQSSSLDSSSTSSQHLNHSKSVTPASTLTGSGP-GRWKTPAIPAT 1020

QY 211 -ISLNSVR-----LKAELMAL-----NYLGIQPT----- 236

DB 1021 PVAVSQPIRTLPPLPPPPPPVHYAGFDGMSMDLPLPPPPSANOIGLPSAQVAAAEERKR 1080

QY 237 -----KEOREALRE-----QVQADSKGTVSFGDFV----- 261

DB 1081 EEHORWYEKEKARLEBEREKREQEKLGQMTQSLNPAPSPPLTAQMKPEKSTLQR 1140

QY 262 -----QVARSFLCLQDEVNVGVHETIPSTLDSQLPDCPSLEADVGKLRQERN 309







```

QY 126 SEDCG-----PQSTSTPILLSSPSETILLPKTSSPTQDSS-----159
DB 780 SDRGSGKPRPKSEGFELYNSTONGSPSPQLPWAIESEPKKLPGDRLMKRNRADHRSS 839
QY 160 -----TF 161
DB 840 PNVAQNPPSPGKSAVASGTTAKITSVSTGNLCTEETPPRPEAVPIPTQTYTREYTF 899
QY 162 PSCKA-----IQTKPE-HDKTEHSPI-----181
DB 900 PASKSQDRMAPPONMNYEKEPHMTDSNHSSIAIQRTVSQEELREDKAYQLERHRIE 959
QY 182 -----TSLDSPA-----DTSNADIAPAWTDDSDGPOCK-----210
DB 960 AAMDRKSDSDMWINQSSSLDSSSQEHLNHSKSVTPASTLTGSGP-GRWKTPAIPAT 1018
QY 211 -ISLNPVS-----LKAEXLEMAL-----NYLGIQPT-----236
DB 1019 PVAVSQPIRTDLPPPPPPVHVAGDFGMSMDLPLPPPPSANQIGLPSAQVAABERRKR 1078
QY 237 -----KEOREALRE-----OVQADSKGTVSFGDFV-----261
DB 1079 BEHQRYEKEKARLEERERKREQRKLCQMTQSLNPAPSPPLTAQOMKEKPESTLQR 1138
QY 262 -----QVARSFLCLQDEVNMGVHEIPISILDSQLLPCDSLEADEVGKLRQERN 309
DB 1139 PQETVIRELQPOQPTIERRDLQYITVSKELSS--GTVCPPDPKWRDAKEKLEKQOQ 1195
QY 310 AALEERNVLEKLESKHKQLIEBONVQEKAKAVASETRALRSI-----HLAEA 362
DB 1196 MHIVD-----MLSKIEQLQSKPDRSAESDRLRKLMLEWQFKRLQES 1239
QY 363 AQROAHQEMDYEE-----VIRLLEAEVSELKAO-----LADYSDQNKESVQD 405
DB 1240 KOKD-----EDJEEEDDDVTMLIMQRLAEARRARTAMPALSVLDLLQDEERRRQOQLEE 1295
QY 406 LRKRVTVLDCQLKSKSEMAKFAKASTER 433
DB 1296 MKKREA--EDRARQEEERREERTKR 1321

RESULT 11
Q9NU90
ID Q9NU90 PRELIMINARY; PRT; 1296 AA.
AC Q9NU90;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE DJ470824.1.1 (Myeloid/lymphoid or mixed-lineage leukemia (Trithorax
DE (Drosophila) homolog), translocated to, 4 (AF-6) (Isoform 1))
DE (Fragment).
GN MLLT4.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Williams S.;
RL Submitted (FEB-2000) to the EMBL/GenBank/DBJ databases.
CC -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL; AL049698; CAB76852.1; -.
DR HSSP; Q12923; 3PDZ
DR InterPro; IPR002710; DIL.
DR InterPro; IPR000253; FHA.
DR InterPro; IPR001478; PDZ.
DR Pfam; PF01843; DIL; 1.
DR Pfam; PF00498; FHA; 1.
DR Pfam; PF00595; PDZ; 1.
DR SMART; SM00240; FHA; 1.
DR SMART; SM00228; PDZ; 1.
DR PROSITE; PS50106; PDZ; 1.

```

```

FT NON TER 1 1
SQ SEQUENCE 1296 AA; 145808 MW; 9PECCF3CA67D19EC CRC64;
Query Match 6.3%; Score 178; DB 4; Length 1296;
Best Local Similarity 18.5%; Pred. No. 0.043;
Matches 110; Conservative 94; Mismatches 194; Indels 198; Gaps 21;
QY 11 SSPLDRDPAPRVITVTKETGLGKILGINRNEGPI-VYIHEVPGGDCYKGRKLKPGQD 69
DB 666 AQPLREPIIITVLKKQMGSLVAAKAGCQDKLGIYKSVVKGGAADVDRLAAGDQ 725
QY 70 LVSINKESIGVSFEBAKSIITRAKLRSSEWIEAFIRKQS-YCGHPGNICPSP--QV 125
DB 726 LLSVDGRSLVGLSQERAAELMTKTS--SVVTLEVA--KQAIYHGLATLLNQSPMWQRI 781
QY 126 SEDCG-----PQSTSTPILLSSPSETILLPKTSSPTQ-----155
DB 782 SDRGSGKPRPKSEGFELYNSTONGSPSPQLPWAIESEPKKLPGDRLMKRNRADHRSS 841
QY 156 -----TQDSTFPSCAKIQTKEPHDKTEHSPTISLDSN-----187
DB 842 PNVAQNPPSPGKSAVASGTTAKITSVSTGNLCTEETPPRPEAVPIPTQTYTREYTF 901
QY 188 PAKTSNADIAP---AWTDDSGP-----OGKISLNPVSRLKAEKLE 225
DB 902 PASKSQDRMAPPONMNYEKEPHMTDSNHSSIAIQRTVSQEELREDKAYQLERHRIE 961
QY 226 MALN-----YLGIOPTKEQREALREQVQADSKGTVS-----256
DB 962 AAMDRKSDSDMWINQSSSLDSSSQEHLNHSKSVTPASTLTGSGPGRWKTPAIPATP 1021
QY 257 -----FGDFVQVARSFLCLQDEVN-VGVHEIPISILDSQLLPC 293
DB 1022 VAVSQPIRTDLPPPPPPVHVAGDFGMSMDLPLPPPPSANQIG--LPS---AQVAAA 1075
QY 294 DSLEAEVGLKROERNAALEERNVLK-----EKL 322
DB 1076 ERKRBEHQRYEKEKARLEERERKREQRKLCQMTQSLNPAPSPPLTAQOMKEKPEP 1335
QY 323 LESEKHKQLIEBONVQEKAKAVASETRALR-----SRIHLAEAAQROAHQEMDYEE-- 376
DB 1136 STLQRPQETVIRELQ-PQOQPTI--ERRDLQYITVSKELSSGDSLSLDPKWRDAKEKL 1192
QY 377 -----VIRLLEAEVSELKAOQLADYSDQNKESVQDLRKRVTVLDCQLKSKSEMARK 425
DB 1193 EKQOMQHIVDMLSKIEQLQSK-----PDRSAESDRLRKLMLEWQFKRLQESKOK 1244

RESULT 12
Q35890
ID Q35890 PRELIMINARY; PRT; 1663 AA.
AC Q35890;
DT 01-JAN-1998 (TrEMBLrel. 05, Created)
DT 01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE S-Afadin.
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RA MEDLINE=97477397; PubMed=9334353;
RA Mandai K., Nakanishi H., Satoh A., Obaishi H., Wada M., Nishioka H.,
RA Itoh M., Mizoguchi A., Aoki T., Fujimoto T., Matsuda Y., Tsukita S.,
RA Takai Y.;
RT "Afadin: A novel actin filament-binding protein with one PDZ domain
RT localized at cadherin-based cell-to-cell adherens junction.";
RL J. Cell Biol. 139:517-528(1997)
CC -!- SIMILARITY: CONTAINS 1 PDZ/DHR DOMAIN.
DR EMBL; U83231; AAC53391.1; -.
DR HSSP; O12923; 3PDZ.
DR InterPro; IPR002710; DIL.

```







GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:59:20 ; Search time 282.013 Seconds  
(without alignments)  
1797.171 Million cell updates/sec

Title: US-09-647-978A-2  
Perfect score: 2832  
Sequence: 1 MSDCTASRSSPLDRDPAF.....LSCAESEDPCRELTDPKS 557

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues  
Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Pending Patents AA Main:\*

1:	/cgn2_6/ptodata/1/paa/US06_COMB.pcp.*
2:	/cgn2_6/ptodata/1/paa/US07_COMB.pcp.*
3:	/cgn2_6/ptodata/1/paa/US08_COMB.pcp.*
4:	/cgn2_6/ptodata/1/paa/US09_COMB.pcp.*
5:	/cgn2_6/ptodata/1/paa/US081_COMB.pcp.*
6:	/cgn2_6/ptodata/1/paa/US082_COMB.pcp.*
7:	/cgn2_6/ptodata/1/paa/US083_COMB.pcp.*
8:	/cgn2_6/ptodata/1/paa/US084_COMB.pcp.*
9:	/cgn2_6/ptodata/1/paa/US085_COMB.pcp.*
10:	/cgn2_6/ptodata/1/paa/US086_COMB.pcp.*
11:	/cgn2_6/ptodata/1/paa/US087_COMB.pcp.*
12:	/cgn2_6/ptodata/1/paa/US088_COMB.pcp.*
13:	/cgn2_6/ptodata/1/paa/US089_COMB.pcp.*
14:	/cgn2_6/ptodata/1/paa/US090_COMB.pcp.*
15:	/cgn2_6/ptodata/1/paa/US091_COMB.pcp.*
16:	/cgn2_6/ptodata/1/paa/US092_COMB.pcp.*
17:	/cgn2_6/ptodata/1/paa/US093_COMB.pcp.*
18:	/cgn2_6/ptodata/1/paa/US094_COMB.pcp.*
19:	/cgn2_6/ptodata/1/paa/US095_COMB.pcp.*
20:	/cgn2_6/ptodata/1/paa/US096_COMB.pcp.*
21:	/cgn2_6/ptodata/1/paa/US097A_COMB.pcp.*
22:	/cgn2_6/ptodata/1/paa/US097B_COMB.pcp.*
23:	/cgn2_6/ptodata/1/paa/US098_COMB.pcp.*
24:	/cgn2_6/ptodata/1/paa/US099A_COMB.pcp.*
25:	/cgn2_6/ptodata/1/paa/US099B_COMB.pcp.*
26:	/cgn2_6/ptodata/1/paa/US100_COMB.pcp.*
27:	/cgn2_6/ptodata/1/paa/US101_COMB.pcp.*
28:	/cgn2_6/ptodata/1/paa/US102_COMB.pcp.*
29:	/cgn2_6/ptodata/1/paa/US103_COMB.pcp.*
30:	/cgn2_6/ptodata/1/paa/US104_COMB.pcp.*
31:	/cgn2_6/ptodata/1/paa/US106_COMB.pcp.*
32:	/cgn2_6/ptodata/1/paa/US106_COMB.pcp.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2832	100.0	557	20	US-09-647-978A-2

2	2832	100.0	557	22	US-09-791-537-144266	Sequence 144266,
3	2109	74.5	553	20	US-09-647-978A-5	Sequence 5, Appli
4	1345	47.5	405	1	PCT-US02-22833-13	Sequence 13, Appl
5	1345	47.5	405	1	PCT-US02-22833A-13	Sequence 13, Appl
6	1193	42.1	300	26	US-10-012-697-1551	Sequence 1551, Ap
7	1193	42.1	300	26	US-10-012-697-1551	Sequence 1551, Ap
8	639.5	22.6	256	32	US-60-236-804-455	Sequence 455, App
9	574.5	20.3	217	1	PCT-US01-47004-181	Sequence 181, App
10	574.5	20.3	217	30	US-10-416-991-181	Sequence 181, App
11	412	14.5	150	1	PCT-US01-01329-1354	Sequence 1354, Ap
12	412	14.5	150	1	PCT-US01-01329-1354	Sequence 1354, Ap
13	412	14.5	150	22	US-09-764-891-3931	Sequence 3931, Ap
14	412	14.5	150	26	US-10-080-090-1354	Sequence 1354, Ap
15	412	14.5	150	28	US-10-205-303-1354	Sequence 1354, Ap
16	195.5	6.3	1492	22	US-09-791-537-23557	Sequence 23557, A
17	188.5	6.7	549	1	PCT-US01-08631-33775	Sequence 33775, A
18	187.5	6.6	1427	22	US-09-791-537-23553	Sequence 23553, A
19	187.5	6.6	1500	22	US-09-791-537-23553	Sequence 23560, A
20	187.5	6.6	1743	1	PCT-US01-08631-41287	Sequence 41287, A
21	187.5	6.6	1743	1	PCT-US01-08656-9308	Sequence 9308, Ap
22	187.5	6.6	1743	28	US-10-273-573-9308	Sequence 9308, Ap
23	187.5	6.6	1816	28	US-10-219-051B-2337	Sequence 2337, Ap
24	187.5	6.6	1816	28	US-10-219-051B-2341	Sequence 2341, Ap
25	187.5	6.6	1816	28	US-10-219-051B-4973	Sequence 4973, Ap
26	187.5	6.6	1816	28	US-10-219-051B-13391	Sequence 13391, A
27	187.5	6.6	1816	28	US-10-219-051B-14543	Sequence 14543, A
28	187.5	6.6	1816	30	US-10-408-765A-1352	Sequence 1352, Ap
29	187.5	6.6	1816	32	US-60-389-987-1352	Sequence 1352, Ap
30	187.5	6.6	1816	32	US-60-412-418-1352	Sequence 1352, Ap
31	184.5	6.5	549	1	PCT-US01-08631-42643	Sequence 42643, A
32	183.5	6.5	1333	22	US-09-791-537-23563	Sequence 23563, A
33	182.5	6.4	762	30	US-10-437-963-122020	Sequence 122020,
34	182.5	6.4	883	26	US-10-094-749-2707	Sequence 2707, Ap
35	181.5	6.4	1410	22	US-09-791-537-53271	Sequence 53271, A
36	181.5	6.4	1410	30	US-10-408-765A-1199	Sequence 1199, Ap
37	181.5	6.4	1410	32	US-60-389-987-1199	Sequence 1199, Ap
38	181.5	6.4	1410	32	US-60-412-418-1199	Sequence 1199, Ap
39	181	6.4	1178	1	PCT-US02-13142-8240	Sequence 8240, Ap
40	181	6.4	1178	27	US-10-128-714-8240	Sequence 8240, Ap
41	178	6.3	1296	22	US-09-791-537-23564	Sequence 23564, A
42	178	6.3	1612	9	US-08-545-860C-48	Sequence 48, Appl
43	178	6.3	1612	9	US-08-545-860C-48	Sequence 48, Appl
44	178	6.3	1612	10	US-08-686-059-2	Sequence 2, Appli
45	177.5	6.3	935	22	US-09-757-781-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1  
US-09-647-978A-2  
; Sequence 2, Application US/09647978A  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; TITLE OF INVENTION: Gene Encoding Syntaxin Interacting Protein  
; FILE REFERENCE: 09/647,978  
; CURRENT APPLICATION NUMBER: US/09/647,978A  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: PCT/US99/08568  
; PRIOR FILING DATE: 1999-04-19  
; PRIOR APPLICATION NUMBER: 60/082,454  
; PRIOR FILING DATE: 1998-04-20  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 2  
; LENGTH: 557  
; TYPE: PRT  
; ORGANISM: Mouse  
; US-09-647-978A-2

Query Match 100.0%; Score 2832; DB 20; Length 557;  
Best Local Similarity 100.0%; Pred. No. 1.1e-207;  
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
QY 1 MSDGTASARSSPLDRDPAPFRVITVTKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
DB 1 MSDGTASARSSPLDRDPAPFRVITVTKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
QY 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
DB 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
QY 121 PPSQVSEDCGQPTSTFTLLSSPSETLLPKTSSTPQTQDSTFPSCAKIQTKEPHDKTEHSP 180
DB 121 PPSQVSEDCGQPTSTFTLLSSPSETLLPKTSSTPQTQDSTFPSCAKIQTKEPHDKTEHSP 180
QY 181 ITSLSNPSADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQR 240
DB 181 ITSLSNPSADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQR 240
QY 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLQDEVNNGVHVEIPSIILDSQLLPCDSLEADE 300
DB 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLQDEVNNGVHVEIPSIILDSQLLPCDSLEADE 300
QY 301 VVKLRQERNAALBERNVLEKLESEKHKRQKLIIEELQNVKQAKAVAEETRALRSRIHLA 360
DB 301 VVKLRQERNAALBERNVLEKLESEKHKRQKLIIEELQNVKQAKAVAEETRALRSRIHLA 360
QY 361 EAAQROAHGEMDMYEEVIRLLAEVSELKAQADYSDQNKESVQDLRKRVTVLDCQLRKS 420
DB 361 EAAQROAHGEMDMYEEVIRLLAEVSELKAQADYSDQNKESVQDLRKRVTVLDCQLRKS 420
QY 421 EMARKAFKASTERLLGFIETAEIQEVLDDSSAPLSTLSERRAVLASQTSPLLARNGRSPPA 480
DB 421 EMARKAFKASTERLLGFIETAEIQEVLDDSSAPLSTLSERRAVLASQTSPLLARNGRSPPA 480
QY 481 TLLLESKELVRSVRAILDMDCLPYGWEAYTADGKYFINHVTQTTSMIHPVMSALNLSC 540
DB 481 TLLLESKELVRSVRAILDMDCLPYGWEAYTADGKYFINHVTQTTSMIHPVMSALNLSC 540
QY 541 AESEEDCPRELTDPKS 557
DB 541 AESEEDCPRELTDPKS 557
RESULT 2
US-09-791-537-144266
; Sequence 144266, Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bionomix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danzer, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAMILY MEMBE
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 144266
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-791-537-144266
Query Match 100.0%; Score 2832; DB 22; Length 557;
Best Local Similarity 100.0%; Pred. No. 1.1e-207;
Matches 557; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MSDGTASARSSPLDRDPAPFRVITVTKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
DB 1 MSDGTASARSSPLDRDPAPFRVITVTKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
QY 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
DB 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
```

```
QY 121 PPSQVSEDCGQPTSTFTLLSSPSETLLPKTSSTPQTQDSTFPSCAKIQTKEPHDKTEHSP 180
DB 121 PPSQVSEDCGQPTSTFTLLSSPSETLLPKTSSTPQTQDSTFPSCAKIQTKEPHDKTEHSP 180
QY 181 ITSLSNPSADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQR 240
DB 181 ITSLSNPSADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQR 240
QY 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLQDEVNNGVHVEIPSIILDSQLLPCDSLEADE 300
DB 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLQDEVNNGVHVEIPSIILDSQLLPCDSLEADE 300
QY 301 VVKLRQERNAALBERNVLEKLESEKHKRQKLIIEELQNVKQAKAVAEETRALRSRIHLA 360
DB 301 VVKLRQERNAALBERNVLEKLESEKHKRQKLIIEELQNVKQAKAVAEETRALRSRIHLA 360
QY 361 EAAQROAHGEMDMYEEVIRLLAEVSELKAQADYSDQNKESVQDLRKRVTVLDCQLRKS 420
DB 361 EAAQROAHGEMDMYEEVIRLLAEVSELKAQADYSDQNKESVQDLRKRVTVLDCQLRKS 420
QY 421 EMARKAFKASTERLLGFIETAEIQEVLDDSSAPLSTLSERRAVLASQTSPLLARNGRSPPA 480
DB 421 EMARKAFKASTERLLGFIETAEIQEVLDDSSAPLSTLSERRAVLASQTSPLLARNGRSPPA 480
QY 481 TLLLESKELVRSVRAILDMDCLPYGWEAYTADGKYFINHVTQTTSMIHPVMSALNLSC 540
DB 481 TLLLESKELVRSVRAILDMDCLPYGWEAYTADGKYFINHVTQTTSMIHPVMSALNLSC 540
QY 541 AESEEDCPRELTDPKS 557
DB 541 AESEEDCPRELTDPKS 557
RESULT 3
US-09-647-978A-5
; Sequence 5, Application US/09647978A
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: Gene Encoding Syntaxin Interacting Protein
; FILE REFERENCE: 09/647,978
; CURRENT APPLICATION NUMBER: US/09/647,978A
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/US99/08568
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: 60/082,454
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 5
; LENGTH: 553
; TYPE: PRT
; ORGANISM: Human
US-09-647-978A-5
Query Match 74.5%; Score 2109; DB 20; Length 553;
Best Local Similarity 76.5%; Pred. No. 3e-152;
Matches 427; Conservative 49; Mismatches 76; Indels 6; Gaps 4;
QY 1 MSDGTASARSSPLDRDPAPFRVITVTKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
DB 1 MNKNTSTVWSPSLLEKDPAFOMITIAKETGLGLKILGGINRNEGPLVYIHEVIPGGDCYK 60
QY 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
DB 61 DGRLLPGDQVLSINKESMIGVSFEAKSIITRAKLRSSESPWEIAFIRKSYCGHPGNICC 120
QY 121 PPSQVSEDCGQPTSTFTLLSSPSETLLPKTSSTPQTQDSTFPSCAKIQTKEPHDKTEHSP 179
DB 121 TSLIEASGEYGPQASTLFSFSPPEILIPKTSSTPTKTNNDILSSC---EIKCYNKTVOI 177
QY 180 PITSLDNSPADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQ 239
DB 180 PITSLDNSPADTSNADIAPAWTDDSGPGQKISLNPSVRLKAEKLEMAINYLGIOPTKEQ 239
```

Db 178 PITS-ENSTVGLSNTDVASAWT-ENYGLQEKISUNPSVRFKAEKLEMALNYLGIQPTKEQ 235  
Qy 240 REALREOVQADSKGTVSFGDFVQVARSFLCQLDDEVNNGVHEIPSLDLSQLLPCDSLEAD 299  
Db 236 HQALRQOVQADSKGTVSFGDFVQVARNLFLCQLDDEVNNGVHEIPSLDLSQLLPCDSSEAD 295  
Qy 300 EVGKLROERNAALFEERNVLKELLESSEKHKRLIEBLQNVKQKAVAEETRALRSRIHL 359  
Db 296 EMERLKCERDALKVNTLLEKLESQKQKQTEBLQNVKQKAVAEETRALRSRIHL 355  
Qy 360 AEAARQAHGEMDYEIVIRLLEAEVSELKAQALADYSQNKESVQDLRKRVTVLDCQLRK 419  
Db 356 AEAARQAHGEMDYEIVIRLLEAEVSELKAQALADYSQNKESVQDLKKRIMVLDQQLRK 415  
Qy 420 SEMARKAKASTERLGLFIRAIQEVLDSSAPLTSERRAVLASOTSPLLRNRSRSP 479  
Db 416 SEMARKTFEASTKELHFBAIQEVFSDNTPISNLSERRAVLASOTSPLLRNRSRSP 475  
Qy 480 ATLLLESKELVRSVRALDMDCLPYGWEEAYTAGIKYFINHYVTTOTSMIHPVMSALNLS 539  
Db 476 ATLALESKELVRSVRALDMDCLPYGWEBAYTAGIKYFINHYVTTOTSMIHPVMSVJNLS 535  
Qy 540 CABESEEDCPRELTDPKS 557  
Db 536 RSENEEDCSRELDPNOKS 553

## RESULT 4

PCT-US02-22833-13

; Sequence 13, Application PC/TUS0222833

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: LAL, Preeti G.

; APPLICANT: HONNCHHELL, Cynthia D.

; APPLICANT: FORSYTHE, Ian J.

; APPLICANT: WALIA, Narinder K.

; APPLICANT: TANG, Y. Tom

; APPLICANT: BOROMSKI, Mark L.

; APPLICANT: BARROSO, Ines

; APPLICANT: WARREN, Bridget A.

; APPLICANT: THANGAVELU, Kavitha

; APPLICANT: GIETZEN, Kimberly J.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LEE, Ernestine A.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: GORVAD, Ann E.

; APPLICANT: DUGGAN, Brendan M.

; APPLICANT: TRAN, Bao

; APPLICANT: LI, Joana X.

; APPLICANT: RICHARDSON, Thomas W.

; APPLICANT: ELLIOTT, Vicki S.

; APPLICANT: ZEBARUJADIAN, Yeganeh

; APPLICANT: TRAN, Uyen K.

; APPLICANT: YAO, Monique G.

; APPLICANT: PETERSON, David P.

; APPLICANT: LUO, Wen

; APPLICANT: LEHR-WASON, Patricia M.

; TITLE OF INVENTION: RECEPTORS AND MEMBRANE ASSOCIATED PROTEINS

; FILE REFERENCE: PP-1082 PCT

; CURRENT APPLICATION NUMBER: PCT/US02/22833

; CURRENT FILING DATE: 2002-07-16

; PRIOR FILING DATE: 2002-07-16

; PRIOR APPLICATION NUMBER: US 60/306,020

; PRIOR FILING DATE: 2001-07-17

; PRIOR APPLICATION NUMBER: US 60/308,179

; PRIOR FILING DATE: 2001-07-27

; PRIOR APPLICATION NUMBER: US 60/309,702

; PRIOR FILING DATE: 2001-08-02

; PRIOR APPLICATION NUMBER: US 60/311,476

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/311,718

; PRIOR FILING DATE: 2001-08-10

; PRIOR APPLICATION NUMBER: US 60/311,551

; PRIOR FILING DATE: 2001-08-10  
; PRIOR APPLICATION NUMBER: US 60/314,798  
; PRIOR FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/316,639  
; PRIOR FILING DATE: 2001-08-31  
; PRIOR APPLICATION NUMBER: US 60/317,996  
; PRIOR FILING DATE: 2001-09-07  
; NUMBER OF SEQ ID NOS: 46  
; SOFTWARE: PERL Program  
; SEQ ID NO 13  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; NAME/KEY: misc feature  
; OTHER INFORMATION: Incyte ID No: 2187465CD1  
PCT-US02-22833-13

Query Match 47.5%; Score 1345; DB 1; Length 405;  
Best Local Similarity 69.5%; Pred. No. 7.3e-94;  
Matches 280; Conservative 37; Mismatches 58; Indels 28; Gaps 5;

Qy 1 MSDGTASARSSSPDRDPAPRVITVTKETGLGLKILGGINRNEGPLVYIHEVPGDDCYK 60  
Db 1 MNKNTSTVWSPSLLEKDPAPQMITIAKETGLGLKVLGGINRNEGPLVYIHEVPGDDCYK 60  
Qy 61 DGRLLKPGDQLVSNKESMIGVSPEEAKSIITRAKLRSESPWIAPIROKSYCGHPCNICC 120  
Db 61 DGRLLKPGDQLVSNKESMIGVSPEEAKSIITRAKLRSESPWIAPIROKSDMIQENLSC 120  
Qy 121 PS-POVSEDCGPOTSTFTLLSSPSETLLPKTSSTPOTQDSTFPSCAKIOTKPEHDKTEHS 179  
Db 121 TSLIEASGEYGPQASTLSLFSPPPEILIPKTSSTPKTNNDILSSC---EIKGYNKTVQI 177  
Qy 180 PITSLDMSPADTSDNADIAPAWTDDSGPOGKISLNPVSRLKAEKLEMALNYLGIQPTKEQ 239  
Db 178 PITS-ENSTVGLSNTDVASAWT-ENYGLQEKISLNPVSRLKAEKLEMALNYLGIQPTKEQ 235  
Qy 240 REALREOVQADSKGTVSFGDFVQVARSFLCQLDDEVNNGVHEIPSLDLSQLLPCDSLEAD 299  
Db 236 HQALRQOVQADSKGTVSFGDFVQVARNLFLCQLDDEVNNGVHEIPSLDLSQLLPCDSSEAD 295  
Qy 300 EVGKLROERNAALFEERNVLKELLESSEKHKRLIEBLQNVKQKAVAEETRALRSRIHL 359  
Db 296 EMERLKCERDALKVNTLLEKLESQKQKQTEBLQNVKQKAVAEETRALRSRIHL 355  
Qy 360 AEAARQAHGEMDYEIVIRLLEAEVSELKAQALADYSQNKES 402  
Db 334 AEAARQAHGEMDYEIVIRLLEAEVSELKAQALADYSQNKES 376

## RESULT 5

PCT-US02-22833A-13

; Sequence 13, Application PC/TUS0222833A

; GENERAL INFORMATION:

; APPLICANT: INCYTE GENOMICS, INC.

; APPLICANT: LAL, Preeti G.

; APPLICANT: HONNCHHELL, Cynthia D.

; APPLICANT: FORSYTHE, Ian J.

; APPLICANT: WALIA, Narinder K.

; APPLICANT: TANG, Y. Tom

; APPLICANT: BOROMSKI, Mark L.

; APPLICANT: BARROSO, Ines

; APPLICANT: YUE, Henry

; APPLICANT: WARREN, Bridget A.

; APPLICANT: THANGAVELU, Kavitha

; APPLICANT: GIETZEN, Kimberly J.

; APPLICANT: AZIMZAI, Yalda

; APPLICANT: LEE, Ernestine A.

; APPLICANT: BAUGHN, Mariah R.

; APPLICANT: GORVAD, Ann E.

; APPLICANT: DUGGAN, Brendan M.

; APPLICANT: TRAN, Bao

```

; APPLICANT: LI, Joana X.
; APPLICANT: RICHARDSON, Thomas W.
; APPLICANT: ELLIOTT, Vicki S.
; APPLICANT: ZEBARJADIAN, Veganeh
; APPLICANT: TRAN, Uyen K.
; APPLICANT: YAO, Monique G.
; APPLICANT: PETERSON, David P.
; APPLICANT: LUO, Wen
; APPLICANT: LEHR-MASON, Patricia M.
; TITLE OF INVENTION: RECEPTORS AND MEMBRANE ASSOCIATED PROTEINS
; FILE REFERENCE: PF-1082 PCT
; CURRENT APPLICATION NUMBER: PCT/US02/22833A
; CURRENT FILING DATE: 2002-09-17
; PRIOR APPLICATION NUMBER: US 60/306,020
; PRIOR FILING DATE: 2001-07-17
; PRIOR APPLICATION NUMBER: US 60/308,179
; PRIOR FILING DATE: 2001-07-27
; PRIOR APPLICATION NUMBER: US 60/309,702
; PRIOR FILING DATE: 2001-08-02
; PRIOR APPLICATION NUMBER: US 60/311,476
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/311,718
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/311,551
; PRIOR FILING DATE: 2001-08-10
; PRIOR APPLICATION NUMBER: US 60/314,798
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/316,639
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: US 60/317,996
; PRIOR FILING DATE: 2001-09-07
; NUMBER OF SEQ ID NOS: 46
; SOFTWARE: PERL Program
; SEQ ID NO 13
; LENGTH: 405
; TYPE: PRT
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; OTHER INFORMATION: Incyte ID No: 2187465CD1
PCT-US02-22833A-13

```

```

Query Match 47.5%; Score 1345; DB 1; Length 405;
Best Local Similarity 69.5%; Pred. No. 7.3e-94;
Matches 280; Conservative 37; Mismatches 58; Indels 28; Gaps 5;

QY 1 MSDDTASARSSPLDRDPAFRVITVTKETGLGKILGGINRNEGFLVYIHEVIPGGDCYK 60
DB 1 MNKNTSTVSPSLEKDPAFQMITIAKETGLGKILGGINRNEGFLVYIHEVIPGGDCYK 60

QY 61 DGRLLKPGDLYSINKESMIGVSFEAKSIITRAKLRESPEWIEAFIROKSYCGHPGNC 120
DB 61 DGRLLKPGDLYSINKESMIGVSFEAKSIITRAKLRESPEWIEAFIROKSYCGHPGNC 120

QY 121 PS-PQVSDCCGQPTFTLLSSPSETLLPKTSSPTQDSTFFSCKAIOTKPEHDKTEHS 179
DB 121 TSLIEASGEYGFQASTLSIFSPPELIPKTSPTKNTNDLLSSC---EIKTYNKTQVI 177

QY 180 PITSLDNGSPADTSNADIAPMTDDSGPGQKISLNPSVRLKAEKLEMALNYLGIQPTKEQ 239
DB 178 PITSLDNGSPADTSNADIAPMTDDSGPGQKISLNPSVRLKAEKLEMALNYLGIQPTKEQ 235

QY 240 REALREQVQADSKGTYSFGDFVQVARSFLCLQDENVVGVHEIPSLDLSOLLPCDSEAD 299
DB 236 HOALRQVQVQADSKGTYSFGDFVQVARNFLCLQDENVVGVHEIPSLDLSOLLPCDSEAD 295

QY 300 EVGKLRQERNAALEERNVLEKLESEKRRKQLIEQLQNVQKAEKAVAEETRALRSRIHL 359
DB 296 EMERLKCERDDALKEVNTLK-----EAKAVVEETRALRSRIHL 333

QY 360 AEAQOQAHGEMDMYEEVIRLLEAEVSELKLAQLADYSQNKES 402
DB 334 AEAQOQAHGEMDMYEEVIRLLEAEKITELKLAQLADYSQNKES 376

```

```

RESULT 6
PCT-US02-28214-1551
; Sequence 1551, Application PC/TUS0228214
; GENERAL INFORMATION:
; APPLICANT: Escobedo, Jaime
; APPLICANT: Garcia, Pablo Dominguez
; APPLICANT: Kabsam, Altaf
; APPLICANT: Lamson, George
; APPLICANT: Scott, Beth
; APPLICANT: Drmanac, Radoje
; APPLICANT: Crkvenjakov, Radomir
; APPLICANT: Dickson, Mark
; APPLICANT: Drmanac, Snezana
; APPLICANT: Labat, Ivan
; APPLICANT: Leshkowitz, Dena
; APPLICANT: Kita, David
; APPLICANT: Garcia, Veronica
; APPLICANT: Jones, Lee William
; APPLICANT: Stache-Crain, Birgit
; TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS
; FILE REFERENCE: 2300-16252
; CURRENT APPLICATION NUMBER: PCT/US02/28214
; CURRENT FILING DATE: 2002-09-04
; PRIOR APPLICATION NUMBER: 60/254,648
; PRIOR FILING DATE: 2000-12-07
; PRIOR APPLICATION NUMBER: 60/275,668
; PRIOR FILING DATE: 2001-03-13
; NUMBER OF SEQ ID NOS: 1568
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1551
; LENGTH: 300
; TYPE: PRT
; ORGANISM: Homo sapiens
PCT-US02-28214-1551

Query Match 42.1%; Score 1193; DB 1; Length 300;
Best Local Similarity 51.8%; Pred. No. 2.1e-82;
Matches 261; Conservative 18; Mismatches 21; Indels 204; Gaps 3;

QY 17 DPAFRVITVTKETGLGKILGGINRNEGFLVYIHEVIPGGDCYKDGRLKPGDLYSINK 76
DB 1 DPAFQMITIAKETGLGKILGGINRNEGFLVYIHEVIPGGDCYKDGRLKPGDLYSINK 60

QY 77 SMIGVSFEAKSIITRAKLRESPEWIEAFIROKSYCGHPGNCPSQVSDCCGPTSTF 136
DB 61 SMIGVSFEAKSIITGAKL----- 79

QY 137 TLLSSPSETLLPKTSSPTQDSTFFSCKAIOTKPEHDKTEHSPTSLDNGSPADTSNADI 196
DB 80 ----- 79

QY 197 APMTDDSGPGQKISLNPSVRLKAEKLEMALNYLGIQPTKEQREALREQVQADSKGTYS 256
DB 80 -----SQARKEEKEWALNYLGIQPTKEQHQAALRQVQADSKGTYS 121

QY 257 FGDVQVARSFLCLQDENVVGVHEIPSLDLSOLLPCDSEADEVQKLRQERNAALEERN 316
DB 122 FGDVQVARNFLCLQDENVVGVHEIPSLDLSOLLPCDSEADEVQKLRQERNAALEERN 154

QY 317 VLKELKLESEKRRKQLIEQLQNVQKAEKAVAEETRALRSRIHLAEAAQOAHGEMDMYEE 376
DB 155 ---EKLLESQKRLQTEELQNVQKAEKAVVEETRALRSRIHLAEAAQOAHGEMDMYEE 211

QY 377 VIRLLEAEVSELKLAQLADYSQNKESVQDLRKEVTVLDCQLRKSEMARKAFKASTERLLG 436
DB 212 VIRLEAKITELKLAQLADYSQNK----- 235

QY 437 PIEAIQEVLLDSSAPLSTLERRAVLASQTSPLPARNGRSPFATLLLESKELVRSVRAI 496
DB 236 -----SERBAVLASQTSPLPARNGRSPFATLLLESKELVRSVRAI 276

```

QY 497 LDMDCLPYCHWEEAYTADGIKYFIN 520  
 Db 277 LDMDCLPYCHWEEAYTADGIKYFIN 300

RESULT 7  
 US-10-012-697-1551  
 ; Sequence 1551, Application US/10012697  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Escobedo, Jaime  
 ; APPLICANT: Garcia, Pablo Dominguez  
 ; APPLICANT: Kassam, Altat  
 ; APPLICANT: Lamson, George  
 ; APPLICANT: Scott, Beth  
 ; APPLICANT: Drmanac, Radole  
 ; APPLICANT: Crkvenjakov, Radomir  
 ; APPLICANT: Dickson, Mark  
 ; APPLICANT: Drmanac, Snezana  
 ; APPLICANT: Labat, Ivan  
 ; APPLICANT: Leshkowitz, Dena  
 ; APPLICANT: Kita, David  
 ; APPLICANT: Garcia, Veronica  
 ; APPLICANT: Jones, Lee William  
 ; APPLICANT: Stache-Crain, Birgit  
 ; TITLE OF INVENTION: HUMAN GENES AND GENE EXPRESSION PRODUCTS  
 ; TITLE OF INVENTION: ISOLATED FROM HUMAN PROSTATE  
 ; FILE REFERENCE: 2300-16252  
 ; CURRENT APPLICATION NUMBER: US/10/012.697  
 ; CURRENT FILING DATE: 2003-01-21  
 ; PRIOR APPLICATION NUMBER: 60/254,648  
 ; PRIOR FILING DATE: 2000-12-07  
 ; PRIOR APPLICATION NUMBER: 60/275,668  
 ; PRIOR FILING DATE: 2001-03-13  
 ; NUMBER OF SEQ ID NOS: 1568  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 1551  
 ; LENGTH: 300  
 ; TYPE: PRT  
 ; ORGANISM: Homo sapiens  
 US-10-012-697-1551

Query Match 42.1%; Score 1193; DB 26; Length 300;  
 Best Local Similarity 51.8%; Pred. No. 2.1e-82;  
 Matches 261; Conservative 18; Mismatches 21; Indels 204; Gaps 3;

QY 17 DPAPRVITVETGLGKILGINRNEGPLYVHEVTPGDCYKDGRLKPGDQVLSINKE 76  
 Db 1 DPAPQMITIAKETGLGLVGLGINRNEGPLYVTEIIPGGDCYKDGRLKPGDQVLSVKNK 60

QY 77 SMIGVSFEAKSIITRAKLRSESPWEIAFIRQKSYCGHPGNICPSPQVSEDCGPOTSTF 136  
 Db 61 SMIGVSFEAKSIITGAKL----- 79

QY 137 TLLSPSETLPKTSSTPQTODSTFPSCAIQTAKPEHDKTEHSPITSLONSPADTSNADI 196  
 Db 80 ----- 79

QY 197 APATDDSDSGPGKISLNPVRLKAEKLEMAALNYGIQPTKEOREALREQVQADSKGTVS 256  
 Db 80 -----SQARKKEKEEMALNYGIQPTKEHQALRQVQADSKGTVS 121

QY 257 FGFVQVARSFCLQDDEVNVGVHEIPSLDSQLPCDSLEAEVGVKLRQERNALERN 316  
 Db 122 FGFVQVARNLFCQLDEVNVGAHISILDSQ----- 154

QY 317 VLKELSEKHKRKLIELQNVKQEAQVAETRALRSRIHLAAQAQAHGMEMDYEE 376  
 Db 155 -----EKLSEDKRKQKTEELQNVKQEAQVAETRALRSRIHLAAQAQAHGMEMDYEE 211

QY 377 VIRLLEAEVSELKAQADYSQNKESVQDLRKRVTVLDCQLRKSEMARKAFKASTERLLG 436  
 Db 212 VIRLLEAKITELKAQADYSQNK----- 235

QY 437 FIEAIQEVLLDSSAPLSTLSERRAVLASQTSPLIARNGRSFPATLLLESKELVRSVRAI 496  
 Db 236 -----SERRAVLASQTSPLTPLGRNGRSIPATLALLESKELVRSVRAI 276

QY 497 LDMDCLPYCHWEEAYTADGIKYFIN 520  
 Db 277 LDMDCLPYCHWEEAYTADGIKYFIN 300

RESULT 8  
 US-60-236-804-455  
 ; Sequence 455, Application US/60236804  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Ladunga, Steven  
 ; APPLICANT: Spier, Eugene  
 ; APPLICANT: Greenberg, Simon  
 ; APPLICANT: Rabkin, Steven  
 ; APPLICANT: Wang, Yu  
 ; APPLICANT: Bonazzi, Vivien  
 ; TITLE OF INVENTION: ISOLATED HUMAN SECRETED PROTEINS,  
 ; TITLE OF INVENTION: NUCLEIC ACID MOLECULES ENCODING HUMAN SECRETED PROTEINS, AND  
 ; TITLE OF INVENTION: USES THEREOF  
 ; FILE REFERENCE: CL000866  
 ; CURRENT APPLICATION NUMBER: US/60/236,804  
 ; CURRENT FILING DATE: 2000-10-02  
 ; NUMBER OF SEQ ID NOS: 1071  
 ; SOFTWARE: FastSeq for Windows Version 4.0  
 ; SEQ ID NO 455  
 ; LENGTH: 256  
 ; TYPE: PRT  
 ; ORGANISM: HUMAN  
 US-60-236-804-455

Query Match 22.6%; Score 639.5; DB 32; Length 256;  
 Best Local Similarity 60.2%; Pred. No. 4.6e-40;  
 Matches 150; Conservative 17; Mismatches 33; Indels 49; Gaps 5;

QY 277 VGVHEIPSTLDSQLPCD---SL--EADVGKLRQERNAALERNVLKELSEKHKR 330  
 Db 25 IDTEIPEI--AKLCCSYGSTSLIWEADEMELKCRDDALKVENTLKEKLESQKXK 82

QY 331 QLIEELQNVKQEAQVAETRALRSRIHLAAQAQAHGMEMDYEEVIRLLEAEVSELKA 390  
 Db 83 QLIEELQNVKQEAQVAETRALRSRIHLAAQAQAHGMEMDYEEVIRLLEAKITELKA 142

QY 391 QLADYSDONKESVQDLRKRVTVLDCQLRKSEMARKAFKASTERLJGF:EAICEVLLDSSA 450  
 Db 143 QLADYSDQNK-----AIQEVFSDNST 163

QY 451 PLSTLSERRAVLASQTSPLIARNGRSFPATLLLESKELVRSVRAILDMDCLPYCHWEEAY 510  
 Db 164 PLSNLSERRAVLASQTSPLTPLGRNGRSIPATLALLESKELVRSVRAILDMDCLPYCHWEEAY 510

QY 511 --TADGIKY 517  
 Db 224 DGTLDLVRY 232

RESULT 9  
 PCT-US01-47004-181  
 ; Sequence 181, Application PC/TUS0147004  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Hyseq, Inc  
 ; TITLE OF INVENTION: NOVEL NUCLEIC ACIDS AND POLYPEPTIDES  
 ; FILE REFERENCE: 21272-090  
 ; CURRENT APPLICATION NUMBER: PCT/US01/47004  
 ; CURRENT FILING DATE: 2002-07-22  
 ; PRIOR APPLICATION NUMBER: 09/728,952  
 ; PRIOR FILING DATE: 2000-11-30  
 ; NUMBER OF SEQ ID NOS: 186  
 ; SOFTWARE: PatentIn version 3.0  
 ; SEQ ID NO 181



```

; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PC006
; CURRENT APPLICATION NUMBER: US/09/764,891
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 10231
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 3931
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-891-3931

Query Match          14.5%; Score 412; DB 22; Length 150;
Best Local Similarity 61.4%; Pred. No. 5.5e-23;
Matches 94; Conservative 14; Mismatches 39; Indels 6; Gaps 4;

Cy 79 IGVSFEAKSIITRAKLRSPEWIEAFIROKSYCGHPGNICCPSPQVSEDCGPQTSTFT 137
Db 1 IGVSFEAKSIITRAKLRSPEWIEAFIROKSDNIQPNLSCTSLIEASGEYGPQASTLS 60
Cy 138 LLSPPSETLLPKTSTPTQDSTFPSCAKIOTKPEHDKTEHSPTSLDNPADTSNADIA 197
Db 61 LFSPPPEILIPKTSSTPKTNDILSSC---EIKTGYNKTVQIPITS-ENSTVGLSNTDVA 116
Cy 198 PAWTDSDSGPGKISLNPVRLKAEKLEMALNY 230
Db 117 SAWT-ENYGLQEKISLNPVRFKAEKLEMALNY 148

RESULT 14
US-10-080-090-1354
; Sequence 1354, Application US/10080090
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA120C1
; CURRENT APPLICATION NUMBER: US/10/080,090
; CURRENT FILING DATE: 2002-02-22
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 3506
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1354
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-080-090-1354

Query Match          14.5%; Score 412; DB 26; Length 150;
Best Local Similarity 61.4%; Pred. No. 5.5e-23;
Matches 94; Conservative 14; Mismatches 39; Indels 6; Gaps 4;

Cy 79 IGVSFEAKSIITRAKLRSPEWIEAFIROKSYCGHPGNICCPSPQVSEDCGPQTSTFT 137
Db 1 IGVSFEAKSIITRAKLRSPEWIEAFIROKSDNIQPNLSCTSLIEASGEYGPQASTLS 60
Cy 138 LLSPPSETLLPKTSTPTQDSTFPSCAKIOTKPEHDKTEHSPTSLDNPADTSNADIA 197
Db 61 LFSPPPEILIPKTSSTPKTNDILSSC---EIKTGYNKTVQIPITS-ENSTVGLSNTDVA 116
Cy 198 PAWTDSDSGPGKISLNPVRLKAEKLEMALNY 230
Db 117 SAWT-ENYGLQEKISLNPVRFKAEKLEMALNY 148
```

```

RESULT 15
US-10-205-303-1354
; Sequence 1354, Application US/10205303
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PA120C2
; CURRENT APPLICATION NUMBER: US/10/205,303
; CURRENT FILING DATE: 2002-07-26
; Prior Application removed - See File Wrapper or Palm
; NUMBER OF SEQ ID NOS: 3506
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1354
; LENGTH: 150
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (149)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-10-205-303-1354

Query Match          14.5%; Score 412; DB 28; Length 150;
Best Local Similarity 61.4%; Pred. No. 5.5e-23;
Matches 94; Conservative 14; Mismatches 39; Indels 6; Gaps 4;

Cy 79 IGVSFEAKSIITRAKLRSPEWIEAFIROKSYCGHPGNICCPSPQVSEDCGPQTSTFT 137
Db 1 IGVSFEAKSIITRAKLRSPEWIEAFIROKSDNIQPNLSCTSLIEASGEYGPQASTLS 60
Cy 138 LLSPPSETLLPKTSTPTQDSTFPSCAKIOTKPEHDKTEHSPTSLDNPADTSNADIA 197
Db 61 LFSPPPEILIPKTSSTPKTNDILSSC---EIKTGYNKTVQIPITS-ENSTVGLSNTDVA 116
Cy 198 PAWTDSDSGPGKISLNPVRLKAEKLEMALNY 230
Db 117 SAWT-ENYGLQEKISLNPVRFKAEKLEMALNY 148
```

Search completed: October 9, 2003, 11:11:44  
Job time : 284.013 secs



Result No.	Score	Query Match	Length	DB ID	Description
1	100	100	100	1	1

1	2832	100.0	557	21	AAV52446	Murine syntxin-4
2	2832	100.0	557	23	AAO15046	Mouse syntxin 4 1
3	2000.5	70.6	554	23	AAO15047	Human syntxin 4 1
4	574.5	20.3	217	23	ABG66753	Human novel polype
5	412	14.5	150	22	ABB95970	Human testicular a
6	412	14.5	150	22	AAV95273	Human reproductive
7	243	8.6	72	23	AAO15045	Human syntxin 4 1
8	188.5	6.7	549	22	ABG03416	Novel human diagno
9	187.5	6.6	1743	22	ABG10928	Novel human diagno

## ALIGNMENTS

15

DE Murine syntaxin-4 interacting protein (SYNTP)

Syntaxin-4 interacting protein; SNRP; glucose; transport; GLUT4;  
vesicle translocation; insulin; regulation; SNARE; SNARE-like;  
soluble N-ethylmaleimide-sensitive factor attachment protein receptor  
uptake; syntaxin-4; VAMP2; competition; binding; glucose storage;  
glucose utilization; recombinant expression; gene therapy; diagnostic  
agonist; diabetes; glycogen storage disease; obesity;  
type II; polycystic ovarian syndrome; hypertension; atherosclerosis;  
insulin resistance; antidiabetic; anorectic; hypotensive;  
arteriosclerosis.

PN WO9954465-A2.  
XX 28-OCT-1999.  
XX  
XX  
XX 19-APR-1999; 99WO-US08568  
XX 20-APR-1998; 98US-0082454  
XX (WARN ) WARNER LAMBERT CO.  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX  
XX Min J, Pessin JE, Saltiel AR, Syu L;  
XX WPI; 2000-038498/03.  
XX  
XX Novel polypeptides and polynucleotides used for diagnosis of syndromes  
XX involving abnormal levels of glucose or abnormal GLUT4 translocation -  
XX  
XX Claim 6; Fig 1A; 51pp; English.  
XX  
XX This sequence represents murine syntaxin-4 interacting protein (SYNIP),  
XX which is a novel insulin-regulated SNARE-like protein directly involved  
XX in the regulation of glucose transport and GLUT4 glucose transporter  
XX vesicle translocation. Insulin induces translocation of GLUT4 from the  
XX intracellular low density microsomal compartment to the cell surface,  
XX GLUT4 translocation playing an important role in the uptake of glucose  
XX by cells. Insulin-stimulated glucose transport and GLUT4 translocation  
XX require specific interactions between the vesicle membrane SNARE  
XX (soluble N-ethylmaleimide-sensitive factor attachment protein receptor),  
XX VAMP2, and the target membrane SNARE, syntaxin-4. SYNIPs competitively  
XX bind to syntaxin-4, preventing the ligand from interacting with its  
XX cognate intracellular receptor, and are only expressed in cells which  
XX exhibit insulin-responsive glucose transport and GLUT4 translocation.  
XX Insulin induces a dissociation of the SYNIP:syntaxin-4 complex via a  
XX decrease in the binding affinity of SYNIP for syntaxin-4. Binding of the  
XX SYNIP C-terminal domain is in contrast refractive to insulin stimulation,  
XX but inhibits glucose transport and GLUT4 translocation. SYNIP proteins  
XX and nucleotides may be used in treatment of a variety of disease states  
XX characterised by abnormal GLUT4 translocation or abnormal glucose storage  
XX and/or utilisation. SYNIP nucleotides may be used to recombinantly  
XX express SYNIP proteins, in gene therapy, or as a source of diagnostic  
XX probes and primers. SYNIP proteins may be used to identify antagonists  
XX which will prevent the binding of SYNIP to syntaxin-4, thereby increasing  
XX glucose transport, or agonists, which will act to decrease glucose  
XX transport. The diseases that may be treated include diabetes  
XX (particularly type II), glycogen storage diseases, obesity, polycystic  
XX ovarian syndrome, hypertension, atherosclerosis and other diseases  
XX associated with insulin resistance.  
XX Note: SYNIP cDNAs (mouse and human), and an additional SYNIP protein are  
XX also claimed, but the sequences are not given in the specification.  
XX  
XX Sequence 557 AA;

QY 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLDEVAVGVHEIPISILDSQLPCDSLEADE 300  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 241 EALREQVQADSKGTVSFGDFVQVARSFLCQLDEVAVGVHEIPISILDSQLPCDSLEADE 300  
QY 301 VGKLRORERNAALERNVLKELLESSEKHKQIIEELQNVKQEKAKAEETRAIERSRIHLA 360  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 301 VGKLRORERNAALERNVLKELLESSEKHKQIIEELQNVKQEKAKAEETRAIERSRIHLA 360  
QY 361 EAAQROAHGEMENDYBEVIRLLEAEVSELKAQIADYSDQNKESVQDLKRKRVTLVDCQLRKS 420  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 361 EAAQROAHGEMENDYBEVIRLLEAEVSELKAQIADYSDQNKESVQDLKRKRVTLVDCQLRKS 420  
QY 421 EMARKAFKASTRELIGFTEAQEVLLDSSAPLSTLSERRAVLASQTSPLLRNCRSFPFA 480  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 421 EMARKAFKASTRELIGFTEAQEVLLDSSAPLSTLSERRAVLASQTSPLLRNCRSFPFA 480  
QY 481 TLLESSELVRSVRAILDMDCLPYGWEEAYTADGKIYFINHVTQTTSWIHVMSALNLSLSC 540  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 481 TLLESSELVRSVRAILDMDCLPYGWEEAYTADGKIYFINHVTQTTSWIHVMSALNLSLSC 540  
QY 541 ABESEEDCPRELTDPKS 557  
DB ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||  
XX 541 ABESEEDCPRELTDPKS 557  
RESULT 2  
AAOI5046  
ID AAOI5046 standard; Protein; 557 AA.  
XX AAOI5046;  
AC AAOI5046;  
XX  
DT 16-AUG-2002 (first entry)  
XX  
DE Mouse syntaxin 4 interacting protein.  
XX  
KW Mouse; antisense gene therapy; Syntaxin 4 interacting protein;  
KW antisense oligonucleotide; diabetes; obesity; skeletal muscle disorder;  
KW inflammation; tumour formation.  
XX  
OS Mus musculus.  
XX  
PN WO200224864-A2.  
XX  
XX 28-MAR-2002.  
XX  
XX 19-SEP-2001; 2001WO-US29251.  
XX  
XX 22-SEP-2000; 2000US-0668313.  
XX  
XX (ISIS-) ISIS PHARM INC.  
XX  
PI Monia BP, Freier SM, Wyatt JR;  
XX  
XX WPI; 2002-401986/43.  
DR N-PSDB; ABQ62254.  
XX  
PT Novel antisense compound that hybridizes and inhibits nucleic acid  
PT molecule encoding Syntaxin 4 interacting protein, useful for treating  
XX diabetes, obesity and skeletal muscle disorder -  
XX  
XX Disclosure; Page 95-98; 154pp; English.  
XX  
XX The invention comprises antisense oligonucleotides designed to inhibit  
XX expression of Syntaxin 4 interacting protein. The antisense  
XX oligonucleotides of the invention are useful for inhibiting the  
XX expression of Syntaxin 4 interacting protein in cells or tissues. The  
XX antisense oligonucleotides are also useful for treating an animal having  
XX a disease or condition associated with Syntaxin 4 interacting protein  
XX (e.g. diabetes, obesity or a skeletal muscle disorder). The antisense  
XX oligonucleotides can also be used to prevent or delay infection,  
XX inflammation and tumour formation. The present amino acid sequence  
XX represents a mouse Syntaxin 4 interacting protein.  
XX





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:59:20 , Search time 279.987 Seconds  
(without alignments)  
1797.171 Million cell updates/sec

Title: US-09-647-978A-5

Perfect score: 2784

Sequence: 1 MNKNTSTWGSLLKDPAP.....LSRSENEEDCSRELPNQKS 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 5728757 seqs, 909918778 residues

Total number of hits satisfying chosen parameters: 5728757

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Pending Patents AA Main.\*

- 1: /cgn2\_6/ptodata/1/paa/PCTUS COMB.pcp.\*
- 2: /cgn2\_6/ptodata/1/paa/US06 COMB.pcp.\*
- 3: /cgn2\_6/ptodata/1/paa/US07 COMB.pcp.\*
- 4: /cgn2\_6/ptodata/1/paa/US08 COMB.pcp.\*
- 5: /cgn2\_6/ptodata/1/paa/US081 COMB.pcp.\*
- 6: /cgn2\_6/ptodata/1/paa/US082 COMB.pcp.\*
- 7: /cgn2\_6/ptodata/1/paa/US083 COMB.pcp.\*
- 8: /cgn2\_6/ptodata/1/paa/US084 COMB.pcp.\*
- 9: /cgn2\_6/ptodata/1/paa/US085 COMB.pcp.\*
- 10: /cgn2\_6/ptodata/1/paa/US086 COMB.pcp.\*
- 11: /cgn2\_6/ptodata/1/paa/US087 COMB.pcp.\*
- 12: /cgn2\_6/ptodata/1/paa/US088 COMB.pcp.\*
- 13: /cgn2\_6/ptodata/1/paa/US089 COMB.pcp.\*
- 14: /cgn2\_6/ptodata/1/paa/US090 COMB.pcp.\*
- 15: /cgn2\_6/ptodata/1/paa/US091 COMB.pcp.\*
- 16: /cgn2\_6/ptodata/1/paa/US092 COMB.pcp.\*
- 17: /cgn2\_6/ptodata/1/paa/US093 COMB.pcp.\*
- 18: /cgn2\_6/ptodata/1/paa/US094 COMB.pcp.\*
- 19: /cgn2\_6/ptodata/1/paa/US095 COMB.pcp.\*
- 20: /cgn2\_6/ptodata/1/paa/US096 COMB.pcp.\*
- 21: /cgn2\_6/ptodata/1/paa/US097A COMB.pcp.\*
- 22: /cgn2\_6/ptodata/1/paa/US097B COMB.pcp.\*
- 23: /cgn2\_6/ptodata/1/paa/US098 COMB.pcp.\*
- 24: /cgn2\_6/ptodata/1/paa/US099A COMB.pcp.\*
- 25: /cgn2\_6/ptodata/1/paa/US099B COMB.pcp.\*
- 26: /cgn2\_6/ptodata/1/paa/US100 COMB.pcp.\*
- 27: /cgn2\_6/ptodata/1/paa/US101 COMB.pcp.\*
- 28: /cgn2\_6/ptodata/1/paa/US102 COMB.pcp.\*
- 29: /cgn2\_6/ptodata/1/paa/US103 COMB.pcp.\*
- 30: /cgn2\_6/ptodata/1/paa/US104 COMB.pcp.\*
- 31: /cgn2\_6/ptodata/1/paa/US106 COMB.pcp.\*
- 32: /cgn2\_6/ptodata/1/paa/US60 COMB.pcp.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2784	100.0	553	20 US-09-647-978A-5	Sequence 5, Appli

2	2109	75.8	557	20	US-09-647-978A-2	Sequence 2, Appli
3	2109	75.8	557	22	US-09-791-537-144266	Sequence 144266,
4	1864	67.0	405	1	PCT-US02-22833-13	Sequence 13, Appl
5	1864	67.0	405	1	PCT-US02-22833A-13	Sequence 13, Appl
6	1312	47.1	300	1	PCT-US02-28214-1551	Sequence 1551, Ap
7	1312	47.1	300	26	US-10-012-697-1551	Sequence 1551, Ap
8	866	31.1	217	1	PCT-US01-47004-181	Sequence 181, App
9	866	31.1	217	30	US-10-416-991-181	Sequence 181, App
10	780.5	28.0	256	32	US-09-236-804-455	Sequence 455, App
11	744	26.7	150	1	PCT-US01-01329-1354	Sequence 1354, Ap
12	744	26.7	150	1	PCT-US01-01329-3931	Sequence 3931, Ap
13	744	26.7	150	22	US-09-764-891-3931	Sequence 3931, Ap
14	744	26.7	150	26	US-10-080-090-1354	Sequence 1354, Ap
15	744	26.7	150	28	US-10-205-103-1354	Sequence 1354, Ap
16	313	11.2	549	1	PCT-US01-08631-42643	Sequence 42643, A
17	310	11.1	549	1	PCT-US01-08631-33775	Sequence 33775, A
18	195	7.0	561	13	US-08-967-909-2	Sequence 2, Appli
19	190.5	6.8	1020	27	US-10-179-131-5887	Sequence 5887, Ap
20	190.5	6.8	1881	1	PCT-US02-03987-15590	Sequence 15590, A
21	190.5	6.8	1881	26	US-10-032-585-7646	Sequence 7646, Ap
22	190.5	6.8	1881	26	US-10-072-851-15590	Sequence 15590, A
23	190.5	6.8	1881	32	US-09-314-050-7646	Sequence 7646, Ap
24	188.5	6.8	2722	21	US-09-724-676-91394	Sequence 91394, A
25	188.5	6.8	2722	21	US-09-724-676A-91394	Sequence 91394, A
26	188.5	6.8	2872	21	US-09-724-676-91401	Sequence 91401, A
27	188.5	6.8	2872	21	US-09-724-676A-91401	Sequence 91401, A
28	188.5	6.8	2929	21	US-09-724-676-91386	Sequence 91386, A
29	188.5	6.8	2929	21	US-09-724-676A-91386	Sequence 91386, A
30	188.5	6.8	2997	21	US-09-724-676-91405	Sequence 91405, A
31	188.5	6.8	2997	21	US-09-724-676A-91405	Sequence 91405, A
32	188.5	6.8	3041	21	US-09-724-676-91393	Sequence 91393, A
33	188.5	6.8	3041	21	US-09-724-676A-91393	Sequence 91393, A
34	188.5	6.8	3054	21	US-09-724-676-91391	Sequence 91391, A
35	188.5	6.8	3054	21	US-09-724-676A-91391	Sequence 91391, A
36	188.5	6.8	3080	21	US-09-724-676-91380	Sequence 91380, A
37	188.5	6.8	3080	21	US-09-724-676A-91380	Sequence 91380, A
38	188.5	6.8	3191	21	US-09-724-676-91400	Sequence 91400, A
39	188.5	6.8	3191	21	US-09-724-676A-91400	Sequence 91400, A
40	188.5	6.8	3204	21	US-09-724-676-91399	Sequence 91399, A
41	188.5	6.8	3204	21	US-09-724-676A-91399	Sequence 91399, A
42	188.5	6.8	3248	21	US-09-724-676-91385	Sequence 91385, A
43	188.5	6.8	3248	21	US-09-724-676A-91385	Sequence 91385, A
44	188.5	6.8	3261	21	US-09-724-676-91384	Sequence 91384, A
45	188.5	6.8	3261	21	US-09-724-676A-91384	Sequence 91384, A

## ALIGNMENTS

RESULT 1  
US-09-647-978A-5  
; Sequence 5, Application US/09647978A  
; GENERAL INFORMATION:  
; APPLICANT: Warner-Lambert Company  
; TITLE OF INVENTION: Gene Encoding Syntaxin Interacting Protein.  
; FILE REFERENCE: 09/647,978  
; CURRENT APPLICATION NUMBER: US/09/647,978A  
; PRIOR FILING DATE: 2000-10-06  
; PRIOR APPLICATION NUMBER: PCT/US99/08568  
; PRIOR FILING DATE: 1999-04-19  
; PRIOR APPLICATION NUMBER: 60/082,454  
; PRIOR FILING DATE: 1998-04-20  
; NUMBER OF SEQ ID NOS: 6  
; SOFTWARE: PatentIn Ver. 2.1  
; SEQ ID NO 5  
; LENGTH: 553  
; TYPE: PRT  
; ORGANISM: Human  
US-09-647-978A-5

Query Match 100.0%; Score 2784; DB 20; Length 553;  
Best Local Similarity 100.0%; Pred. No. 4.3e-200;  
Matches 553; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY	1	MKNKNTSVVPSLSLEKDPAFOMITIAKETGLGLKVLGG:NRNEGPLVYIQTETIPGSDCYK	60
Db	1	MKNKNTSVVPSLSLEKDPAFOMITIAKETGLGLKVLGG:NRNEGPLVYIQTETIPGSDCYK	60
QY	61	DGRLLPGDQOLSVNKESIGVSPFEAKSIITRAKLRLSASWEIAPTROKSDNIQENLSC	120
Db	61	DGRLLPGDQOLSVNKESIGVSPFEAKSIITRAKLRLSASWEIAPTROKSDNIQENLSC	120
QY	121	TSLEIASGEYGPQASTLSLFSPPPELLIPKTSPTPTNNNDILSSCEIKTGYNKTVQIPIT	180
Db	121	TSLEIASGEYGPQASTLSLFSPPPELLIPKTSPTPTNNNDILSSCEIKTGYNKTVQIPIT	180
QY	181	SENSVTGJLNTDVASAWTENTGLOEKISLNPVSVFRAEKLUEMALNYLGTOPTKEOQHAR	240
Db	181	SENSVTGJLNTDVASAWTENTGLOEKISLNPVSVFRAEKLUEMALNYLGTOPTKEOQHAR	240
QY	241	QQVQADSKGTVSFGDFVQVARNLFCQLQDEVNVGAHEISNILDSQLPDCDSEADEMERL	300
Db	241	QQVQADSKGTVSFGDFVQVARNLFCQLQDEVNVGAHEISNILDSQLPDCDSEADEMERL	300
QY	301	KCERDDALKEVNTLKEKLLESKDKORKQLTBELONVQOEKAVVEETRALRSRIHLAAQA	360
Db	301	KCERDDALKEVNTLKEKLLESKDKORKQLTBELONVQOEKAVVEETRALRSRIHLAAQA	360
QY	361	ROAHGEMWMDYBEVIRLLEAKITELKAOLADYSQNKESVQDLKKRIMVLDCOLRKSEMAR	420
Db	361	ROAHGEMWMDYBEVIRLLEAKITELKAOLADYSQNKESVQDLKKRIMVLDCOLRKSEMAR	420
QY	421	KTFEASTEKLLHFVEAIOEVSFDSNSTPLSNLSERRAVLASCTSLTPLGRNGRSIPATLAL	480
Db	421	KTFEASTEKLLHFVEAIOEVSFDSNSTPLSNLSERRAVLASCTSLTPLGRNGRSIPATLAL	480
QY	481	ESKELVKSVRALLDMDCULPYGWEAEYTAGDKIFYINHVTOTTTWIIHPVMSVLNLSRSEEN	540
Db	481	ESKELVKSVRALLDMDCULPYGWEAEYTAGDKIFYINHVTOTTTWIIHPVMSVLNLSRSEEN	540
QY	541	REDCSRELPNQKS	553
Db	541	REDCSRELPNQKS	553

```

RESULT 2
US-09-647-978A-2
; Sequence 2, Application US/09647978A
; GENERAL INFORMATION:
; APPLICANT: Warner-Lambert Company
; TITLE OF INVENTION: Gene Encoding Syntaxis Interacting Protein
; FILE REFERENCE: 09/647,978
; CURRENT APPLICATION NUMBER: US/09/647,978A
; CURRENT FILING DATE: 2000-10-06
; PRIOR APPLICATION NUMBER: PCT/US99/08568
; PRIOR FILING DATE: 1999-04-19
; PRIOR APPLICATION NUMBER: 60/082,454
; PRIOR FILING DATE: 1998-04-20
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 2
; LENGTH: 557
; TYPE: PRT
; ORGANISM: Mouse
US-09-647-978A-2

```

```

Query Match      75.8%; Score 2109; DB 20; Length 557;
Best Local Similarity 76.5%; Pred.No. 2.8e-149;
Matches 427; Conservative 49; Mismatches 179; Indels 6; Gaps 4

Qy      1  MNKNTSTVTVSPSLLEKDPAFOWITIAKGTGLGLKVLGINNNEGPLYVQIIEIPGDCYK 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      1  MSDGTASARSSPLDRDPAFRVITYKGTGLGLKILGINNNEGPLYVTHEIIPGDCYK 60
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy      61  DGRLLKPGDGLVSNKESMIGVFSFEAKSIITRAKLRLSAWEIAFIQKSDNIQENLSC 120
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

Db	61	DGRUKPGDQVLVSINKESWIGVSEFEAKSIIITRAKLRSFSEWIEAFITRQKSYSGCHPGNICC	120
Qy	121	TSLEAEAGEYQPOASTLUSLFSSPPEILIPKTSSTPKTNNDILSSC---BIKTGYNKTVQI	177
Db	121	PS-POVSEDCQFQSTSTLTLSSPSETLLPKTSTPTQDSTPFSCKAIQTKPEHKDIHS	179
Qy	178	PITS-ENSTVGLSNTQVASAWT-ENTYGLQEKISLNFSEVFPAEKLEMANLYLGIQPTKEQ	235
Db	180	PITSLDNSPADTNSNADITAPAWTDDSGSQGKISLNFSEVRLKAEKLEMANLYLGIQPTKEQ	239
Qy	236	HOALRQQVQADSKGTVSFGDFVQVARNLFCQLQDEVDNVGHAHETSNLDSQLLPCDSSPAD	295
Db	240	REALREQVQADSKGTVSFGDFVQVARSFLCQLQDEVNVGVEHIPSILDSQLLPCDSSLEAD	299
Qy	296	EMERUKCERDDALKEVNTLKEKLLESDQKQKQITTEELQNVQKQAKAVVEETRALRSRHL	355
Db	300	EVGKURQERNAALEERNVLKEKLESEKHKQKQIEELQNVQKQAKAVAEETRALRSRHL	359
Qy	356	AEAAQORAHGMEYBEYVIRLLEAKITELKAQLADYSDQNKESVDQKKRIMVLDQQLRK	415
Db	360	AEAAQORAHGMEYBEYVIRLLEAEVSELKAQLADYSDQNKESVQDLRKRVTLVDCQLRK	419
Qy	416	SEMAKRTFEASTEKLLHFVRAI QEVESDNSTPLSNLSERRAVLASQTSILTPLRNGRSIP	475
Db	420	SEMAKAFKASTERILGFI RAIQEVLLDSSAPLSTLSERRAVLASQTSPLLARNGRSP	479
Qy	476	ATLALSELKVSVRALLDMDCPLPYGHEEAYTAGGIKYFINHVTTQTSWIHPVMVSLNLS	535
Db	480	ATLLESLEKVRVRAILDMDCPLPYGHEEAYTAGGIKYFINHVTTQTSWIHPVMGALNLS	539
Qy	536	RSENEEDCSRELPNQKS	553
Db	540	CAEBSEEDCPRELTDPKS	557

```

RESULT 3
US-09-791-537-144266
; Sequence 144266 Application US/09791537
; GENERAL INFORMATION:
; APPLICANT: Bloncmix, Inc.
; APPLICANT: Debe, Derek
; APPLICANT: Danser, Joseph
; TITLE OF INVENTION: THREE DIMENSIONAL STRUCTURES OF PROTEIN FAMILIES AND FAN
; TITLE OF INVENTION: METHODS OF USE THEREOF
; FILE REFERENCE: 261/210
; CURRENT APPLICATION NUMBER: US/09/791,537
; CURRENT FILING DATE: 2001-02-22
; NUMBER OF SEQ ID NOS: 153055
; SOFTWARE: Patentin version 3.0
; SEQ ID NO 144266
; LENGTH: 557
; TYPE: prt
; ORGANISM: Mus musculus
US-09-791-537-144266

```

	Query Match	75.8%	Score 2109,	DB 22;	Length 557;
	Best Local Similarity	76.5%;	Pred. No. 2.8e-149;		
	Matches 427;	Conservative 49;	Mismatches 76;	Indels 6;	Gaps 4
Qy	1	MNKNTSTVVSFSLLEKOPAFOMITIAKETGLGLKVLGGINRNEGPIVYQEIIPGGDCYK	60		
Db	1	MSDGTASARSSPLDRDPAPFRIIVTETGLGLKILGGINRNEGPIVYHEVIPGGDCYK	60		
Qy	61	DGRLLKPGQLVSNKESMIGVSPFEAKSIITRAKLALASAWETAFIROKSDNIQPNLSC	120		
Db	61	DGRLLKPGQLVSNKESMIGVSPFEAKSIITRAKLASESPWELAFIROKSGCHPNCIC	120		
Qy	121	TSLEIASEGYGPQASTLSLSSPPEILIPIKTSSTPTKNNDISSC---	177		
Db	121	PS-QPVSEDCGPQSTFTLLSSPSETLLPIKTSSTPTQDSTFPSCAIQTKPEHDKTEHS	179		
Qy	178	PIITS-ENSTVGLSNTDVASAWT-ENYGLQEKIISLNPVSVPFAEKLEMAIYLGIQPTKEQ	235		

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:59:50 ; Search time 45.336 Seconds  
(without alignments)  
1965.413 Million cell updates/sec

Title: US-09-647-978A-5

Perfect score: 2784

Sequence: 1 MNKNTSTVSPSLKNDPAF.....LSRSENEEDCSRELPNQKS 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 600653 seqs, 161128416 residues

Total number of hits satisfying chosen parameters: 600653

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.\*

- 1: /cgn2\_6/ptodata/1/pubpaa/US07\_PUBCOMB.pep.\*
- 2: /cgn2\_6/ptodata/1/pubpaa/PCT\_NEW\_PUB.pep.\*
- 3: /cgn2\_6/ptodata/1/pubpaa/US06\_NEW\_PUB.pep.\*
- 4: /cgn2\_6/ptodata/1/pubpaa/US06\_PUBCOMB.pep.\*
- 5: /cgn2\_6/ptodata/1/pubpaa/US07\_NEW\_PUB.pep.\*
- 6: /cgn2\_6/ptodata/1/pubpaa/PCTUS\_PUBCOMB.pep.\*
- 7: /cgn2\_6/ptodata/1/pubpaa/US08\_NEW\_PUB.pep.\*
- 8: /cgn2\_6/ptodata/1/pubpaa/US08\_PUBCOMB.pep.\*
- 9: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 10: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 11: /cgn2\_6/ptodata/1/pubpaa/US09\_PUBCOMB.pep.\*
- 12: /cgn2\_6/ptodata/1/pubpaa/US09\_NEW\_PUB.pep.\*
- 13: /cgn2\_6/ptodata/1/pubpaa/US10A\_PUBCOMB.pep.\*
- 14: /cgn2\_6/ptodata/1/pubpaa/US10B\_PUBCOMB.pep.\*
- 15: /cgn2\_6/ptodata/1/pubpaa/US10C\_PUBCOMB.pep.\*
- 16: /cgn2\_6/ptodata/1/pubpaa/US10\_NEW\_PUB.pep.\*
- 17: /cgn2\_6/ptodata/1/pubpaa/US60\_NEW\_PUB.pep.\*
- 18: /cgn2\_6/ptodata/1/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	744	26.7	150	11	US-09-764-891-3931
2	190.5	6.8	1881	12	Sequence 3931, Ap
3	180	6.5	1356	10	Sequence 7646, Ap
4	175.5	6.3	935	10	Sequence 2, Appli
5	175	6.3	2485	9	US-09-757-781-1
6	169.5	6.1	1337	10	US-09-802-669-46
7	169	6.1	1001	15	US-09-757-781-62
8	167.5	6.0	1743	12	US-10-128-714-3240
9	167	6.0	2466	12	US-09-882-227-624
10	161	5.8	2383	15	US-10-177-980-12
11	158.5	5.7	1388	15	US-10-082-830-260
12	158	5.7	2125	10	US-10-446-473-82
13	158	5.7	2649	12	US-09-919-172-29
14	157	5.6	1178	15	US-10-205-219-169
15	156.5	5.6	1948	12	US-10-128-714-8240
					Sequence 7611, Ap

16	155.5	5.6	888	11	US-09-893-519A-73	Sequence 73, Appl
17	155	5.6	967	15	US-10-197-666A-90	Sequence 90, Appl
18	155	5.6	1591	15	US-10-197-666A-92	Sequence 92, Appl
19	153.5	5.5	2139	9	US-09-727-384-6	Sequence 6, Appli
20	153.5	5.5	2139	15	US-10-023-219-4	Sequence 4, Appli
21	152.5	5.5	2871	15	US-10-146-473-41	Sequence 41, Appl
22	152.5	5.5	1786	10	US-09-742-096-3	Sequence 3, Appli
23	150	5.4	3899	15	US-10-171-311-4	Sequence 4, Appli
24	150	5.4	3907	15	US-10-171-311-2	Sequence 2, Appli
25	150	5.4	3917	15	US-10-171-311-8	Sequence 8, Appli
26	150	5.4	3925	15	US-10-171-311-6	Sequence 6, Appli
27	149.5	5.4	767	10	US-09-919-497-59	Sequence 59, Appl
28	149.5	5.4	2037	10	US-09-951-401-3	Sequence 3, Appli
29	149.5	5.4	2037	10	US-09-922-101-3	Sequence 3, Appli
30	149.5	5.4	2037	10	US-09-951-402-3	Sequence 3, Appli
31	149	5.4	709	12	US-10-256-250-15	Sequence 15, Appl
32	149	5.4	724	15	US-10-211-962-22	Sequence 22, Appl
33	149	5.4	725	11	US-09-978-309A-47	Sequence 47, Appl
34	149	5.4	725	12	US-10-256-250-14	Sequence 14, Appl
35	148.5	5.3	724	15	US-10-211-962-21	Sequence 21, Appl
36	148	5.3	676	12	US-10-256-250-16	Sequence 16, Appl
37	148	5.3	1711	10	US-09-771-161A-219	Sequence 219, App
38	148	5.3	1711	10	US-09-771-161A-220	Sequence 220, App
39	147	5.3	1940	12	US-09-738-630-99	Sequence 99, Appl
40	146.5	5.3	1641	14	US-10-017-216-5	Sequence 5, Appli
41	146	5.2	1938	15	US-10-171-311-164	Sequence 164, App
42	146	5.2	1945	11	US-09-927-597-2	Sequence 2, Appli
43	146	5.2	1972	15	US-10-171-311-162	Sequence 162, App
44	146	5.2	1979	11	US-09-927-597-4	Sequence 4, Appli
45	146	5.2	2055	14	US-10-017-216-4	Sequence 4, Appli

## ALIGNMENTS

### RESULT 1

US-09-764-891-3931  
; Sequence 3931, Application US/09764891  
; Publication No. US20030077808A1  
; GENERAL INFORMATION:  
; APPLICANT: Rosen et al.  
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies  
; FILE REFERENCE: PC006  
; CURRENT APPLICATION NUMBER: US/09/764,891  
; CURRENT FILING DATE: 2001-01-17  
; Prior application data removed - consult PALM or file wrapper  
; NUMBER OF SEQ ID NOS: 10231  
; SOFTWARE: PatentIn Ver. 2.0  
; SEQ ID NO 3931  
; LENGTH: 150  
; TYPE: PRT  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (149)  
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids  
US-09-764-891-3931

Query Match		26.7%	Score 744;	DB 11;	Length 150;
Best Local Similarity		100.0%	Pred. No. 3.3e-47;		
Matches 148;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	79	IGVSPFEAKSIITRAKRL	LESAWEIAFIROKSDNIQ	PENLSCTSLIEASGYGQASTLS	138
Db	1	IGVSPFEAKSIITRAKRL	LESAWEIAFIROKSDNIQ	PENLSCTSLIEASGYGQASTLS	60
Qy	139	LFSSPPEILIPKTSPTK	NNDLSSCEIKYKNTVOIP	TSENSTVGLNTDVASAWT	198
Db	61	LFSSPPEILIPKTSPTK	NNDLSSCEIKYKNTVOIP	TSENSTVGLNTDVASAWT	120
Qy	199	ENYGLQEKISLNP	SVRFKAEKLEMALNY	226	
Db	121	ENYGLQEKISLNP	SVRFKAEKLEMALNY	148	





GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:57:41 ; Search time 14.9459 Seconds  
(without alignments)  
1565.502 Million cell updates/sec

Title: US-09-647-978A-5

Perfect score: 2784

Sequence: 1 MNKNTSTVVSFLLEKPAF.....LSRSENEEDCSRLPNQKS 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 328717

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA.\*

- 1: /cgn2\_6/prodata/1/aaa/5A.COMB.pep.\*
- 2: /cgn2\_6/prodata/1/aaa/5B.COMB.pep.\*
- 3: /cgn2\_6/prodata/1/aaa/5A.COMB.pep.\*
- 4: /cgn2\_6/prodata/1/aaa/5B.COMB.pep.\*
- 5: /cgn2\_6/prodata/1/aaa/PTUS.COMB.pep.\*
- 6: /cgn2\_6/prodata/1/aaa/backfiles1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	186	6.7	3248	1	US-08-353-700-1
2	186	6.7	3248	5	PCT-US95-16216-1
3	185.5	6.7	2482	1	US-08-328-254-6
4	175	6.3	2485	3	US-09-290-640-46
5	170.5	6.1	2465	2	US-08-596-291-3
6	170.5	6.1	2465	3	US-09-100-804-3
7	168	6.0	2008	4	US-09-091-501B-8
8	168	6.0	3433	3	US-09-091-501B-10
9	167	6.0	976	3	US-09-104-324B-4
10	167	6.0	2466	3	US-09-080-855-12
11	167	6.0	2466	4	US-09-086-076-12
12	167	6.0	2466	5	PCT-US94-09943-2
13	162	5.8	1829	3	US-09-157-420-1
14	159.5	5.7	1388	4	US-09-572-193-2
15	159.5	5.7	1388	4	US-09-723-262-2
16	159.5	5.7	1388	4	US-09-723-219-2
17	158	5.7	1612	3	US-08-545-860D-48
18	158	5.7	1612	5	PCT-US94-04496-48
19	155	5.6	1184	3	US-09-541-782-2
20	155	5.6	1184	4	US-09-723-820-2
21	154.5	5.5	1129	4	US-09-252-991A-29927
22	152	5.5	896	1	US-08-095-737-2
23	152	5.5	896	1	US-08-480-145-2
24	152	5.5	896	2	US-08-477-389-2
25	151	5.4	1057	4	US-09-107-532A-4789
26	150.5	5.4	3656	4	US-09-134-001C-5080
27	150	5.4	1073	3	US-09-541-782-6

Sequence 6, Appli  
Sequence 8, Appli  
Sequence 11, Appli  
Sequence 3, Appli  
Sequence 22, Appli  
Sequence 21, Appli  
Sequence 1, Appli  
Sequence 31, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 12, Appli  
Sequence 3944, Ap  
Sequence 16965, A  
Sequence 4, Appli  
Sequence 4, Appli  
Sequence 4, Appli

## ALIGNMENTS

RESULT 1

US-08-353-700-1

; Sequence 1, Application US/08353700

; Patent No. 559919

; GENERAL INFORMATION:

; APPLICANT: YEN, TIMOTHY J.

; APPLICANT: RATTNER, JEROME B.

; TITLE OF INVENTION: NUCLEIC ACID ENCODING A

; TITLE OF INVENTION: TRANSIENTLY-EXPRESSED KINETOCHORE PROTEIN,

; NUMBER OF SEQUENCES: 4

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: DANN, DORFMAN, HERRELL AND SKILLMAN

; STREET: 1601 MARKET STREET, SUITE 720

; CITY: PHILADELPHIA

; STATE: PA

; COUNTRY: USA

; ZIP: 19103-2307

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent in Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/353.700

; FILING DATE: 09-DEC-1994

; CLASSIFICATION: 435

; ATTORNEY/AGENT INFORMATION:

; NAME: REED, JANET E.

; REGISTRATION NUMBER: 36,252

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: (215) 563-4100

; TELEFAX: (215) 563-4044

; INFORMATION FOR SEQ ID NO: 1:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 3248 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

; HYPOTHETICAL: NO

; ANTI-SENSE: NO

; ORIGINAL SOURCE:

; ORGANISM: HUMAN

; US-08-353-700-1

Query Match 6.7%; Score 186; DB 1; Length 3248;

Best Local Similarity 21.6%; Pred.No.3.9e-06;

Matches 104; Conservative 80; Mismatches 180; Indels 118; Gaps 19;



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 9, 2003, 10:57:36 ; Search time 40.3541 Seconds  
(without alignments)  
2175.142 Million cell updates/sec

Title: US-09-647-978A-5

Perfect score: 2784

Sequence: 1 MNKNTSTVSVSLLEKPAF.....LSRSENEEDCSRELPNQKS 553

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 1107863

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A\_Geneseq\_19Jun03.\*  
1: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1980.DAT.\*  
2: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1981.DAT.\*  
3: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1982.DAT.\*  
4: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1983.DAT.\*  
5: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1984.DAT.\*  
6: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1985.DAT.\*  
7: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1986.DAT.\*  
8: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1987.DAT.\*  
9: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1988.DAT.\*  
10: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1989.DAT.\*  
11: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1990.DAT.\*  
12: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1991.DAT.\*  
13: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1992.DAT.\*  
14: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1993.DAT.\*  
15: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1994.DAT.\*  
16: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1995.DAT.\*  
17: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1996.DAT.\*  
18: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1997.DAT.\*  
19: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1998.DAT.\*  
20: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA1999.DAT.\*  
21: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2000.DAT.\*  
22: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2001.DAT.\*  
23: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2002.DAT.\*  
24: /SIDSL1/gcgdata/geneseq/geneseq-emb1/AA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2549.5	91.6	554	AAO15047	Human syntaxin 4 i
2	2109	75.8	557	AAV52446	Murine syntaxin-4
3	2109	75.8	557	AAO15046	Mouse syntaxin 4 i
4	866	31.1	217	ABG66753	Human novel polype
5	744	26.7	150	ABBS9570	Human testicular a
6	744	26.7	150	ABBS9570	Human reproductive
7	313	11.2	549	ABG12284	Novel human diagno
8	310	11.1	549	ABG03416	Novel human diagno
9	287	10.3	72	AAO15045	Human syntaxin 4 i

10	195	7.0	561	19	AAW63043	Streptococcus uber
11	190.5	6.8	1881	23	ABP73809	Candida albicans e
12	188.5	6.8	3685	10	AAV90290	Human Duchenne mus
13	187.5	6.7	801	23	ABP68963	Human polyptide
14	187.5	6.7	1384	24	ABP55413	Human MDDT-22 prot
15	187.5	6.7	1404	24	ABP55393	Human MDDT-2 prote
16	187.5	6.7	3685	10	AAV90373	Sequence encoded b
17	186	6.7	3248	17	AAV95795	Kinetochores protei
18	185.5	6.7	2482	16	AAV72826	Human mitotin. Ho
19	185.5	6.7	2482	19	AAV23996	Human mitotin amin
20	185.5	6.7	3210	24	ABU07438	Protein differenti
21	180	6.5	1341	23	AAU99747	Human ASIP-related
22	179.5	6.4	1690	22	ABG61144	Drosophila melanog
23	179.5	6.4	1690	22	ABG61173	Drosophila melanog
24	178	6.4	1327	23	ABBS7163	Mouse ischaemic co
25	175.5	6.3	935	23	AAU99746	Human ASIP-related
26	175	6.3	2485	21	AAV19343	Amino acid sequenc
27	174	6.2	1095	20	AAV80359	An F-actin-combine
28	169.5	6.1	1310	21	AAV59238	A rod shortened dy
29	169.5	6.1	1337	23	AAU99748	Rat ASIP protein (
30	169	6.1	1001	24	ABJ25582	Aspergillus fumiga
31	168	6.0	2008	18	AAV22016	Utrrophin truncated
32	168	6.0	2013	22	ABBS7964	Amino acid sequenc
33	168	6.0	3433	18	AAV22017	Utrrophin. Homo ea
34	167.5	6.0	1743	19	AAV98879	H. pylori GHPO 175
35	167	6.0	976	22	AAV65681	Human SCP-1 muteti
36	167	6.0	976	24	ABP74709	Human SCP-1 protei
37	167	6.0	2466	16	AAV71498	Human protein tyro
38	167	6.0	2466	19	AAV75999	Intracellular prot
39	167	6.0	2466	21	AAV90272	Human PTP1 phosph
40	166.5	6.0	980	24	ABU03527	Angiogenesis-assoc
41	166	6.0	795	23	ABP77430	Human tumour marke
42	166	6.0	1824	19	AAV20300	Human microtubule
43	165	5.9	817	24	AAE33672	Human structural a
44	163	5.9	865	22	AAV95798	Human protein sequ
45	163	5.9	865	22	AAV67610	Amino acid sequenc

## ALIGNMENTS

RESULT 1  
AAO15047  
ID AAO15047 standard; Protein; 554 AA.  
XX AAO15047;  
AC AAO15047;  
DT 16-AUG-2002 (first entry)  
XX Human syntaxin 4 interacting protein 2.  
DE Human; antisense gene therapy; Syntaxin 4 interacting protein;  
KW antisense oligonucleotide; diabetes; obesity; skeletal muscle disorder;  
XX inflammation; tumour formation.  
XX Homo sapiens.  
OS  
XX  
XX WC200224864-A2.  
XX  
XX 28-MAR-2002.  
XX  
XX 19-SEP-2001; 2001WO-US29251.  
XX  
XX 22-SEP-2000; 2000US-0668313.  
XX (ISIS-) ISIS PHARM INC.  
XX Monia BP, Freier SM, Wyatt JR;  
XX WPI; 2002-401986/43.  
XX N-PSDB; ABQ62261.  
XX  
XX Novel antisense compound that hybridizes and inhibits nucleic acid

PT molecule encoding Syntaxin 4 interacting protein, useful for treating  
PT diabetes, obesity and skeletal muscle disorder -  
PS Disclosure; Page 100-104; 154pp; English.  
XX The invention comprises antisense oligonucleotides designed to inhibit  
CC expression of Syntaxin 4 interacting protein. The antisense  
CC oligonucleotides of the invention are useful for inhibiting the  
CC expression of Syntaxin 4 interacting protein in cells or tissues. The  
CC antisense oligonucleotides are also useful for treating an animal having  
CC a disease or condition associated with Syntaxin 4 interacting protein  
CC (e.g. diabetes, obesity or a skeletal muscle disorder). The antisense  
CC oligonucleotides can also be used to prevent or delay infection,  
CC inflammation and tumor formation. The present amino acid sequence  
CC represents a human Syntaxin 4 interacting protein.  
XX  
SQ Sequence 554 AA;  
Query Match 91.6%; Score 2549.5; DB 23; Length 554;  
Best Local Similarity 93.3%; Pred. No. 1.4e-190;  
Matches 519; Conservative 5; Mismatches 27; Indels 5; Gaps 2;  
QY 1 MNKNTSTVWSPSLLEKDPAFQMTIAKETGLGKVLGGINRNEGPIVYQEIIPGDCYK 60  
DB 1 MNKNTSTVWSPSLLEKDPAFQMTIAKETGLGKVLGGINRNEGPIVYQEIIPGDCYK 60  
QY 61 DGRLLPGDQLVSNKESMIGVSFEAKSIITRAKLRESAWETAFIROKSDNIQENLSC 120  
DB 61 DGRLLPGDQLVSNKESMIGVSFEAKSIITRAKLRESAWETAFIROKSDNIQENLSC 118  
QY 121 TSLIASEGVEQFQASTLSLSPPEIL---IPKTSSTPKTNNDILSSCEIKTYGNTVQI 177  
DB 119 VMYITYRSFRRIWTSLSNLIKSFLLLLKYIPKPSPPKTNNDILSSCEIKTYGNTVQI 178  
QY 178 PITSENSTVGLSNTDVASANTENYGLQKISLNPVSFRFAEKLEMANLYGLQPTKEHQ 237  
DB 179 PITSENSTVGLSNTDVASANTENYGLQKISLNPVSFRFAEKLEMANLYGLQPTKEHQ 238  
QY 238 ALRQVQADSKGTVSFGDFVQVARNLFCQLQDEVNVAHEISNLDQLPCDSSEADEM 297  
DB 239 ALRQVQADSKGTVSFGDFVQVARNLFCQLQDEVNVAHEISNLDQLPCDSSEADEM 298  
QY 298 ERLCERDDALKENTLKEKLESDDKORKOLTRELONVQOEAKVVEETRALRSIHAE 357  
DB 299 ERLCERDDALKENTLKEKLESDDKORKOLTRELONVQOEAKVVEETRALRSIHAE 358  
QY 358 AAQRAHGMEMDYEVIRLEAKITELKAQADYSQNKESVDLKKRIMVLDCQLRKSE 417  
DB 359 AAQRAHGMEMDYEVIRLEAKITELKAQADYSQNKESVDLKKRIMVLDCQLRKSE 418  
QY 418 MARXTEASTKLLHFVEATQEVFSDNSTPLSNLSERRAVLASQTSITPLGRNGRSIPAT 477  
DB 419 MARXTEASTKLLHFVEATQEVFSDNSTPLSNLSERRAVLASQTSITPLGRNGRSIPAT 478  
QY 478 LALSKELVKSVALDMDCPLPYGWEAYTAGDKYFINKHVTOTTSWIHPVMSVLNLSRS 537  
DB 479 LALSKELVKSVALDMDCPLPYGWEAYTAGDKYFINKHVTOTTSWIHPVMSVLNLSRS 538  
QY 538 EENEEDCSRELPNQKS 553  
DB 539 EENEEDCSRELPNQKS 554  
RESULT 2  
AAV52446  
ID AAV52446 standard; protein; 557 AA.  
XX  
AC AAV52446;  
XX  
DT 22-FEB-2000 (first entry)  
XX  
TX Murine syntaxin-4 interacting protein (SYNIP).  
XX

KW Syntaxin-4 interacting protein; SYNIP; glucose; transport; GLUT4;  
KW vesicle translocation; insulin; regulation; SNARE; SNARE-like;  
KW soluble N-ethylmaleimide-sensitive factor attachment protein receptor;  
KW uptake; syntaxin-4; VAMP2; competition; binding; glucose storage;  
KW glucose utilization; recombinant expression; gene therapy; diagnostic;  
KW antagonist; agonist; diabetes; glycogen storage disease; obesity;  
KW type II; polycystic ovarian syndrome; hypertension; atherosclerosis;  
KW insulin resistance; antidiabetic; anorectic; hypotensive;  
KW antiarteriosclerotic.  
XX  
OS Mus sp.  
XX  
KH Key Location/Qualifiers  
FH Domain 19..93 "PDZ domain"  
FT Domain 219..264  
FT Domain /note= "Calcium-binding EF-hand domain"  
FT Domain 300..410  
FT Domain /note= "Tandem coiled coil domains"  
FT Domain 500..533  
FT Domain /note= "WW domain"  
XX  
PN W0954465-A2.  
XX  
XX 28-OCT-1999.  
XX  
XX 19-APR-1999; 99WO-US08568.  
XX  
XX 20-APR-1998; 98US-0082454.  
XX  
XX (WARN ) WARNER LAMBERT CO.  
XX (IOWA ) UNIV IOWA RES FOUND.  
XX  
XX Min J, Pessin JE, Saltiel AR, Syu L;  
XX WPI; 2000-038498/03.  
XX  
XX Novel polypeptides and polynucleotides used for diagnosis of syndromes  
XX involving abnormal levels of glucose or abnormal GLUT4 translocation -  
XX Claim 6; Fig 1A; 51pp; English.  
XX  
XX This sequence represents murine syntaxin-4 interacting protein (SYNIP),  
XX which is a novel insulin-regulated SNARE-like protein directly involved  
XX in the regulation of glucose transport and GLUT4 glucose transporter  
XX vesicle translocation. Insulin induces translocation of GLUT4 from the  
XX intracellular low density microsomal compartment to the cell surface,  
XX GLUT4 translocation playing an important role in the uptake of glucose  
XX by cells. Insulin-stimulated glucose transport and GLUT4 translocation  
XX require specific interactions between the vesicle membrane SNARE  
XX (soluble N-ethylmaleimide-sensitive factor attachment protein receptor),  
XX VAMP2, and the target membrane SNARE, syntaxin-4. SYNIPs competitively  
XX bind to syntaxin-4, preventing the ligand from interacting with its  
XX cognate intracellular receptor, and are only expressed in cells which  
XX exhibit insulin-responsive glucose transport and GLUT4 translocation.  
XX Insulin induces a dissociation of the SYNIP:syntaxin-4 complex via a  
XX decrease in the binding affinity of SYNIP for syntaxin-4. Binding of the  
XX SYNIP C-terminal domain is in contrast refractive to insulin stimulation,  
XX but inhibits glucose transport and GLUT4 translocation. SYNIP proteins  
XX and nucleotides may be used in treatment of a variety of disease states  
XX characterised by abnormal GLUT4 translocation or abnormal glucose storage  
XX and/or utilisation. SYNIP nucleotides may be used to recombinantly  
XX express SYNIP proteins, in gene therapy, or as a source of diagnostic  
XX probes and primers. SYNIP proteins may be used to identify antagonists  
XX which will prevent the binding of SYNIP to syntaxin-4, thereby increasing  
XX glucose transport, or agonists, which will act to decrease glucose  
XX transport. The diseases that may be treated include diabetes  
XX (particularly type II), glycogen storage diseases, obesity, polycystic  
XX ovarian syndrome, hypertension, atherosclerosis and other diseases  
XX associated with insulin resistance.  
XX Note: SYNIP cDNAs (mouse and human), and an additional SYNIP protein are  
XX also claimed, but the sequences are not given in the specification.  
XX